

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 00:43:55 ; Search time 6021 Seconds

(Without alignments)
11334.675 Million cell updates/sec

Title: AF133659

Perfect score: 2345
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table: IDENTITY NTC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GeneBdb:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*
29: em_vl:.*
30: em_hcg_hum:.*
31: em_hcg_inv:.*
32: em_hcg_other:.*
33: em_hcg_mus:.*
34: em_hcg_pln:.*
35: em_hcg_rod:.*
36: em_hcg_mam:.*
37: em_hcg_vit:.*
38: em_sy:.*
39: em_higo_hum:.*
40: em_higo_mus:.*
41: em_higo_other:.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2345	100.0	2345	9 AF133659	AF133659 Homo sapi
2	2340.2	99.8	2416	9 AB005289	AB005289 Homo sapi
3	2332	99.4	2369	9 BC006323	BC006323 Homo sapi
4	2328.2	99.3	2333	9 AF078777	AF078777 Homo sapi
5	2248	95.9	2407	6 AF028561	AF028561 Sequence
6	2248	95.9	2407	6 AR100662	AR100662 Sequence
7	2150.4	91.7	2193	9 AK001418	AK001418 Homo sapi
8	1833	78.2	2684	10 MMU43892	MMU43892 Mus musculu
9	564.8	24.1	2979	3 AY051556	AY051556 Drosophila
10	543.6	23.2	2187	8 AF287699	AF287699 Arabidops
11	543.6	23.2	2373	8 AF360334	AF360334 Arabidops
12	543.4	23.2	2037	8 AF287697	AF287697 Arabidops
13	540.4	23.0	2452	8 ATH272202	ATH272202 Arabidops
14	508.6	21.7	2043	8 AF287698	AF287698 Arabidops
15	480.6	20.5	50701	2 AC018231	AC018231 Drosophila
16	480.6	20.5	16682	3 AC093192	AC093192 Drosophila
17	480.6	20.5	172175	3 AC010038	AC010038 Drosophila
18	480.6	20.5	17583	3 AC105264	AC105264 Drosophila
19	480.6	20.5	257867	3 AC005557	AC005557 Drosophila
20	480.6	20.5	303191	3 AE003472	AE003472 Drosophila
21	425.2	18.1	2262	8 SCMDYGENE	X82612 S. cerevisia
22	425.2	18.1	2286	8 SCMDYGENE	X81715 S. cerevisia
23	425.2	18.1	29286	8 SC9952X	Z49212 S. cerevisia
24	418.8	17.9	282610	1 RPXX01	AJ235270 Rickettsi
25	411.4	17.5	11621	1 AE013719	AE013719 Yersinia
26	411.4	17.5	313050	1 AJ414152	AJ414152 Yersinia
27	406	17.3	43227	2 TSP17	AL049917 Arabidops
28	406	17.3	197655	8 ATCHRIV69	AL161573 Arabidops
29	404.6	17.3	36310	8 SPAC8C9	Z99168 S. pombe chr
30	361	15.4	4481	3 AF466308	AF466308 Dictyoste
31	358.6	15.3	5369	6 A94856	A94856 Sequence 3
32	347.8	14.8	43702	8 SPAC15A10	Z97208 S. pombe chr
33	346.4	14.8	10086	1 AE009585	AE009585 Brucella
34	335	14.3	9945	1 AE008036	AE008036 Agrobacte
35	335	14.3	10322	1 AE009070	AE009070 Agrobacte
36	310.4	13.2	346547	1 AP003012	AP003012 Mesorhizo
37	308	13.1	2946	10 BC006634	BC006634 Mus muscu
38	304.4	13.0	391	9 F241872S16	AF241887 Homo sapi
39	304.4	13.0	100839	2 AC002417	AC002417 Homo sapi
40	304.4	13.0	127243	9 AL359545	AL359545 Human DNA
41	304.4	13.0	156799	2 AL359548	AL359548 Homo sapi
42	304	13.0	2406	3 AF497513	AF497513 Caenorhab
43	302.8	12.9	299350	1 SME591786	AL591786 Sinorhizo
44	301.2	12.8	3008	3 AY084209	AY084209 Drosophila
45	292	12.5	2848	10 RNABJ3004	AJ003004 Rattus no

ALIGNMENTS

RESULT 1
AF133659
LOCUS
DEFINITION
Homo sapiens ATP-binding cassette 7 iron transporter (ABC7) mRNA,
complete cds; nuclear gene for mitochondrial product.

ACCESSION
AF133659
VERSION
AF133659.1 GI:49277189

KEYWORDS
ORGANISM

Source
Homo sapiens

REFERENCE

1 (bases 1 to 2345)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Altkmetz, R., Raekind, W.H., Hutchinson, A., Schneck, N.D., Dean, M.
and Koeller, D.M.

TITLE Mutation of a putative mitochondrial iron transporter gene (ABC7) in x-linked sideroblastic anemia and ataxia (XLSA/A)

JOURNAL Hum. Mol. Genet. 8 (5), 743-749 (1999)

MEDLINE 99214014

PUBMED 10196363

REFERENCE 2 (bases 1 to 2345)

AUTHORS Dean, M., Allikmets, R. and Hutchinson, A.

TITLE Direct Submission

JOURNAL Submitted (08-MAR-1999) Human Genetics, National Cancer Institute, NCI-FCRDC, Frederick, MD 21703, USA

FEATURES

Source

1..2345

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xq13"

1..2345

/gene="ABC7"

1..2259

/gene="ABC7"

/product="ATP-binding cassette 7 iron transporter"

/protein_id="AAD33045.1"

/db_xref="GI:4927190"

/translation="MALAMHSMRMAAAAFKRRHSALIRPLVSVSGSPQMRPH
QLGALGTARAYQIDESLSITMORLKGNSQPLDAKALVWPLIEKRTCHGAGG
GLHTDPKEGLKDVDRKIKAMLSYVWKDPPDRARVAISLGIKAKANNIVPDM
PKVADSLNMGSMNLSDAPNTVATATATVILIGVSRAGAAPPENVRVAIVGKVA
ONSIRIRAKNFIHLNIDLFPHLSRQGLSKAIDRETRISTVLSALVNNLPIPH
EVMVLSGLYIKCAQAFALVTLGLGTATFVAVTRKTRFRIEMNDADNANAI
DSLNLNETVKYFNNEERYBAORYDGLKTEPFAKSTSTLMLNFGSALIFSVGLTAI
MLASOGIVAGTLIVGLDMVNGLFLQLSLPNEGLVRETRQALIDMNTLPIILKY
DTQIKDKMASPLQTPOTVAFPDNVHFEXIEGQKVLISFEVPAGKVAIVGSG
SGKSTVILFRFPEPOKSIYLAGONIDVLSLRVAVGVPODAVLFNNTIYNL
LXGNISAPREVVAIVAKLAGHDAIILMPHXYDPVQVGRKISGEGEQVATIRALI
KDPVYLIDEATSSLDSTIBETITIGAMKDVAKHRTSTFIARLSITVDADETIVDDG
KVAERGHGLANPHSLIYSEMHPTQSSRVONHNDPKWEAKENISKEERKLUOEI
VNSVGCNCNCSC"

BASE COUNT 690 a 459 c 565 g 631 t

ORIGIN

Query Match 100.0%; Score 2345; DB 9; Length 2345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCTGCTGGCATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAAG 60

DB 1 ATGGCGCTGCTGGCATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAAG 60

QY 61 CGCCGCACTCCGCAATTCGATCCGGCCTTGTAGTCTGTAGCGGCTCAGGTCGCGAG 120

DB 61 CGCCGCACTCCGCAATTCGATCCGGCCTTGTAGTCTGTAGCGGCTCAGGTCGCGAG 120

QY 121 TGGAGGCCACATCAACTCGCGGCTTGGGAACCGCTCGAGCTTACCAATTCGAGATCA 180

DB 121 TGGAGGCCACATCAACTCGCGGCTTGGGAACCGCTCGAGCTTACCAATTCGAGATCA 180

QY 121 TGGAGGCCACATCAACTCGCGGCTTGGGAACCGCTCGAGCTTACCAATTCGAGATCA 180

DB 121 TGGAGGCCACATCAACTCGCGGCTTGGGAACCGCTCGAGCTTACCAATTCGAGATCA 180

QY 181 TTTAAAGATATACATGCGAGAGATTGGAAAAGCAATTCGAGACAGTTCTTAATGCT 240

DB 181 TTTAAAGATATACATGCGAGAGATTGGAAAAGCAATTCGAGACAGTTCTTAATGCT 240

QY 241 GCAAAGGCTCTCCAGTATGCGCACTGATAGAAAAGAGACATGTTGGCATGTCATGCA 300

DB 241 GCAAAGGCTCTCCAGTATGCGCACTGATAGAAAAGAGACATGTTGGCATGTCATGCA 300

QY 301 GGAGGAGACTCCACACAGACCAGAAAGGTTAAAGATGTTGATCTCGAAAATC 360

DB 301 GGAGGAGACTCCACACAGACCAGAAAGGTTAAAGATGTTGATCTCGAAAATC 360

QY 361 ATTAAGCAATGCTTTTATATGTGTGCCCCAAGACAGCCGATCTAGACTAGAGTT 420

DB 361 ATTAAGCAATGCTTTTATATGTGTGCCCCAAGACAGCCGATCTAGACTAGAGTT 420

QY 421 GCCATTTGCTGGGATTTTGGTGTGTGCAAAAGCCATGAATATGTGTTCCCTCATG 480

DB 421 GCCATTTGCTGGGATTTTGGTGTGTGCAAAAGCCATGAATATGTGTTCCCTCATG 480

QY 481 TTTAAATATGCTGTAGACAGGCTCAACAGATGTGCGGAAACATGCTGAACCTAGAT 540

DB 481 TTTAAATATGCTGTAGACAGGCTCAACAGATGTGCGGAAACATGCTGAACCTAGAT 540

QY 541 GCACCAATATACAGTTGCAACCATGCGCAACAGCACTTCTGATGGCTATGATCAAGA 600

DB 541 GCACCAATATACAGTTGCAACCATGCGCAACAGCACTTCTGATGGCTATGATCAAGA 600

QY 601 GCTGAGGCTGCTTTTAAACGAAGTTGAAATCAGATTTGGCAAGTAGCCAGAT 660

DB 601 GCTGAGGCTGCTTTTAAACGAAGTTGAAATCAGATTTGGCAAGTAGCCAGAT 660

QY 661 TCAATCCGAAGATAGCCAAAATGTCTTCCATCTTCACAACTCGATCTGAGGTTT 720

DB 661 TCAATCCGAAGATAGCCAAAATGTCTTCCATCTTCACAACTCGATCTGAGGTTT 720

QY 721 CACCTGAGCAGACAGACGGGAGCTTATCTAAGGCTATTGACAGAGAACAAAGGATATC 780

DB 721 CACCTGAGCAGACAGACGGGAGCTTATCTAAGGCTATTGACAGAGAACAAAGGATATC 780

QY 781 AGTTTGTCTGAGTGTCTTGGTATTAATCTTCCATCATGTTTGAAGTATGCTT 840

DB 781 AGTTTGTCTGAGTGTCTTGGTATTAATCTTCCATCATGTTTGAAGTATGCTT 840

QY 841 GTCAATGCTGTTTGTATTAACAAATGCGGTGCCAGTTTGTGTTGTTGTTGTTGTTGTT 900

DB 841 GTCAATGCTGTTTGTATTAACAAATGCGGTGCCAGTTTGTGTTGTTGTTGTTGTTGTT 900

QY 901 CTTGGTATCATACAGCATTCACAGTTGACAGTCAACGCTGAGAACTAGATTAGATA 960

DB 901 CTTGGTATCATACAGCATTCACAGTTGACAGTCAACGCTGAGAACTAGATTAGATA 960

QY 961 GAAATGAACAAGAGATATGATGACAGTATGCTGCTATAGACTCACTGCTGAATAT 1020

DB 961 GAAATGAACAAGAGATATGATGACAGTATGCTGCTATAGACTCACTGCTGAATAT 1020

QY 1021 GAACTGTGAAGTATTTTAATGAAGAAATGAACAGACAGATATGATGATTTTGG 1080

DB 1021 GAACTGTGAAGTATTTTAATGAAGAAATGAACAGACAGATATGATGATTTTGG 1080

QY 1081 AAGAGTATGAGACTGCTTCAATGAAAAGTCTCTACTGCTGCTATAGCTGAACTTTGGT 1140

DB 1081 AAGAGTATGAGACTGCTTCAATGAAAAGTCTCTACTGCTGCTATAGCTGAACTTTGGT 1140

QY 1141 CAAAGTCTATTTTCACTGTGCTTTAAACAGCTATTAATGCTGCTGCGCAGTCAGGAAAT 1200

DB 1141 CAAAGTCTATTTTCACTGTGCTTTAAACAGCTATTAATGCTGCTGCGCAGTCAGGAAAT 1200

QY 1201 GTGGCAGTATCCCTTAATCTGTGAGATCTAGTAATGTTGAATGSACTGCTTTTCACTT 1260

DB 1201 GTGGCAGTATCCCTTAATCTGTGAGATCTAGTAATGTTGAATGSACTGCTTTTCACTT 1260

QY 1261 TCAATACCCCTGAACCTTCTGGAAGCTGATATGAGAGACTAGACAGCACTCATAGAT 1320

DB 1261 TCAATACCCCTGAACCTTCTGGAAGCTGATATGAGAGACTAGACAGCACTCATAGAT 1320

QY 1321 ATGAACACCTTGTACTCTACTCAAGGTAGACCCCAATTAAGAACAAGTATGCA 1380

DB 1321 ATGAACACCTTGTACTCTACTCAAGGTAGACCCCAATTAAGAACAAGTATGCA 1380

QY 1381 TCTCCCTTCAATCAACCAAGACAGTACCTGCTGCTTGAATATGTCATTTTGA 1440

DB 1381 TCTCCCTTCAATCAACCAAGACAGTACCTGCTGCTTGAATATGTCATTTTGA 1440

QY 1441 TACATGAGGCGCAAGAAAGTCTTAGTGAATATCTTTGAAGTCCCTGACAGAAAGAA 1500

DB 1441 TACATGAGGCGCAAGAAAGTCTTAGTGAATATCTTTGAAGTCCCTGACAGAAAGAA 1500

QY 1501 GTGGCATTTGAGAGGATGAGGTCAAGGAAAGACAAATATGAGGCTATTAATTTGCGC 1560

```

Db 1501 GTGGCATTGAGAGGAGTGGGTCAAGGAAAGACAAATAGTAGGCTATTTATTCGC 1560
Qy 1561 TTCTATGAGCCTCAAAAGGGTAGCAATTTATCTTGCTGTCACAAATATATCAAGATGTGAGC 1620
Db 1561 TTCTATGAGCCTCAAAAGGGTAGCAATTTATCTTGCTGTCACAAATATATCAAGATGTGAGC 1620
Qy 1621 CTGGAAGCCTTCGAGAGGCGAGTGGAGTGGTACCTCAGATGCTGTCCTCTTCATPAT 1680
Db 1621 CTGGAAGCCTTCGAGAGGCGAGTGGAGTGGTACCTCAGATGCTGTCCTCTTCATPAT 1680
Qy 1681 ACTATTTATTAACAACCTCTTATATGAAACATCAAGTGTCTTACCTGAGGAAGTGTATGCA 1740
Db 1681 ACTATTTATTAACAACCTCTTATATGAAACATCAAGTGTCTTACCTGAGGAAGTGTATGCA 1740
Qy 1741 GTGGCAAAATATAGCTGAGCTTCATGATGCAATCTTCGAATGCCACATGATATGACACC 1800
Db 1741 GTGGCAAAATATAGCTGAGCTTCATGATGCAATCTTCGAATGCCACATGATATGACACC 1800
Qy 1801 CAAGTAGGGGAAACGAGACTCAAGCTTTACAGAGAGAGAAAGAAAGATGAGCAATTGCA 1860
Db 1801 CAAGTAGGGGAAACGAGACTCAAGCTTTACAGAGAGAGAAAGAAAGATGAGCAATTGCA 1860
Qy 1861 AGAGCCATTTTGAAGGACCCCGCAGTCACTCTATGATGAGAGCTACTTATCGTTAGAT 1920
Db 1861 AGAGCCATTTTGAAGGACCCCGCAGTCACTCTATGATGAGAGCTACTTATCGTTAGAT 1920
Qy 1921 TCGATTAAGTGAAGACTATTTCTGTCATGAGAGATGTGTCAAAACAGAGAACTTCT 1980
Db 1921 TCGATTAAGTGAAGACTATTTCTGTCATGAGAGATGTGTCAAAACAGAGAACTTCT 1980
Qy 1981 ATTTTCATTTGCAACAGATTTGTCAACAGTGTGTATGATGATGATGATGATGATGAT 2040
Db 1981 ATTTTCATTTGCAACAGATTTGTCAACAGTGTGTATGATGATGATGATGATGATGAT 2040
Qy 2041 CAGGTAAGGTAAGGCGGAACCGGTGTAACCAACATGTTTCTGTTAACCCTCAATGATC 2100
Db 2041 CAGGTAAGGTAAGGCGGAACCGGTGTAACCAACATGTTTCTGTTAACCCTCAATGATC 2100
Qy 2101 TATTCAGAAATGTGGCATACACAGAGAGCCGTGTGAGAACATGATGATGATGATGATGAT 2160
Db 2101 TATTCAGAAATGTGGCATACACAGAGAGCCGTGTGAGAACATGATGATGATGATGATGAT 2160
Qy 2161 GAAGCAAAAGAAAGAAATATATCCAAAGAGAGAGAAAGAAAGAAAGAAAGAAAT 2220
Db 2161 GAAGCAAAAGAAAGAAATATATCCAAAGAGAGAGAAAGAAAGAAAGAAAGAAAT 2220
Qy 2221 GTCAATAGTGTGAAGAGCTGTGGAACCTGTTCTGTCTAAGTACATTAAGACATTTTCTT 2280
Db 2221 GTCAATAGTGTGAAGAGCTGTGGAACCTGTTCTGTCTAAGTACATTAAGACATTTTCTT 2280
Qy 2281 TTTTGTGTTTGGATACATATTTGCACTGAAGCAGAAATGTTTATTAATAAAATCAT 2340
Db 2281 TTTTGTGTTTGGATACATATTTGCACTGAAGCAGAAATGTTTATTAATAAAATCAT 2340
Qy 2341 ACATT 2345
Db 2341 ACATT 2345

RESULT 2
LOCUS AB005289 2416 bp mRNA linear PRI 16-JUN-1998
DEFINITION Homo sapiens mRNA for ABC transporter 7 protein, complete cde.
ACCESSION AB005289
VERSION AB005289.1 GI:3228278
KEYWORDS ABC transporter 7 protein.
SOURCE Homo sapiens Placenta cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases) Shimada,Y., Okuno,S., Kawai,A., Shinomiya,H., Saito,A., Suzuki,M.,

```

```

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (127-JUN-1997) Yoshikazu Shimada, Otsuka Pharmaceutical
Co. Ltd., Otsuka GSN Research Institute, Kagasuno, Kawachi-cho,
Tokushima, Tokushima 771-0192, Japan
(E-mail:shim@otsuka.genome.ad.jp, Tel:81-886-65-2888,
Fax:81-886-37-1035)
FEATURES
source
1..2416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/lissue_type="placenta"
1..2416
/gene="hABC7"
69..2327
/gene="hABC7"
/codon_start=1
/product="ABC transporter 7 protein"
/protein_id="BAA28861.1"
/db_xref="GI:3228279"
/translation="MALLMASHRWAAAAAFKKRHSAILRPIVSVSGSPQMPH
OUGALGTARAYIPESLSKITRQRLKNGSGFLDRAKLQVPLERKTCCHGAGG
GATDPKGLKQVDRKILKAMLSVWPDRDLARVALISGLFAGAAKVAIVFEM
FYAADVLSQMSGNMLNLSDAENTVAATVAVALIGVSRAGAAFEVRAVNFVFA
ONSIRIADNVFLHNLDLGFLSRQGLSKAIDRGKISFVLSALVNLPIFMF
EMLVSGVLYVCAQFALVLTGATYAFVAVTRMGTREIEMNKANDAGNAAI
DSLNVETKYVNNRERYAORVDFLKTYETSLKSTSLAMNCGSAIPEVGLAI
MYLAAQGIAGTLITVGDVAVNGLLPOLSLPLNGTIVRETRQALIDMTLFTLKV
DTQIKDQWASPLQITPQATVAIFQNVHEVYLEGQVLSISFEVPAKRAKAIVEGSG
SKSTIVRLLEFVEPQKGSIVLACQNIQDVLESIRAVGVVPODAVLFNHTIYVNL
LYGNISASPEEYVAVALAGLDAILRMHGYDQVGERGLKSGEKKORVAIAAIL
KRPVILVDHANSISDIETETILGAMGVVGHRTSIFLAHLSITVDADEITVLDQ
KVAERSTHGLLANPHSITSEMHITQSSRVQNHNDPKWEAKENISKEERKKLOBET
VNSVGGCGNCS"
BASE COUNT 713 a 476 c 577 g 650 t
ORIGIN
Query Match 99.8%; Score 2340.2; DB 9; Length 2416;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGCGCTGCTCGCATGATCTTGGCGCTGGCGCGCGCGCTGTTTCAAAAG 60
Db 69 ATGGCGCTGCTCGCATGATCTTGGCGCTGGCGCGCGCGCTGTTTCAAAAG 128
Qy 61 CGCGCGACTCGCGCATTTGATCGGCTTTAGTCTGTGTAGCGGCTTCAGTCCGAG 120
Db 129 CGCGCGACTCGCGCATTTGATCGGCTTTAGTCTGTGTAGCGGCTTCAGTCCGAG 188
Qy 121 TGGAGGCAATCACTCGCGCGCTTGGGAACCGCTCGAGCCTACCAAGATTCCAGAGTCA 180
Db 189 TGGAGGCAATCACTCGCGCGCTTGGGAACCGCTCGAGCCTACCAAGATTCCAGAGTCA 248
Qy 181 TTAAAAAGTATCATGAGAGATTGGAAAAAGCAATTCAGACAGTCTTAGATGCT 240
Db 249 TTAAAAAGTATCATGAGAGATTGGAAAAAGCAATTCAGACAGTCTTAGATGCT 308
Qy 241 GCAAAGGCTCTCCAGGTATGGCCATGATGAAAGAGAGACATGTTGGCATGTGCA 300
Db 309 GCAAAGGCTCTCCAGGTATGGCCATGATGAAAGAGAGACATGTTGGCATGTGCA 368
Qy 301 GGAAGGAGCTCCACACACCAAGAGGTTTAAAGATGTTGATCTCGGAAATTC 360
Db 369 GGAAGGAGCTCCACACACCAAGAGGTTTAAAGATGTTGATCTCGGAAATTC 428

```


ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2369)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK Contact: MGC help desk
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 DNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/nisc_mgc@nsl.nih.gov
 Contact: nisc_mgc@nsl.nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.U., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantrop, S., Thomas, P.J., Tlonsong, E.B., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 17 Row: k Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
 Location/Qualifiers
 1. 2369
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:12887 IMAGE:4138236"
 /tissue_type="Muscle, Thaddomasarcoma"
 /clone_lib="NIH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: POT87"
 11. 2272
 /codon_start=1
 /product="ATP-binding cassette, sub-family B (MDR/TAP), member 7"
 /translation="MALIMHSRWMAAAAAFEKRHSALIRPIVSVSGSGQWPH QLGALSTARAYQDIPESSKSTIWRKLGSGGFLDARAKLQWPIERKTCHEGAG GSHATPKEGLKQVDTRKI1KAMLSYWPDRDLARVA1SLGFLGAKAMNI1VVF MFKYAVDSINOMGMMNLSDAPNTVAATAV1LIGVSRAGAFAFNEVRN1VFKV AQSIRIRANVFIHLNLDLGHLSRQCALSKALDRGRGSPVLSALVFNUL1IM FEYMLVSVGLYKCGAOPALVTGTTGTAFTVAATPRRTREIRIMNKGAA IDSLNVEYTVKVPNNRXYRQYDFLKYETSLSTSTLANLNGGAS1RVSGLTA IYTLASOGIVAGTLTVGDVWVNGLLF0LSPLNFI1GTYRETR0LLIDMTLFI1LK VDTQIKDKWASPLQITPQTAIVA1FDNVHEVEYLEGQVLSG1SEVPAPAKVA1VGS GSGKSTVRL1FRFEPKQSIYLAGQNTQDVLS1ESL1RAVAVP0DAVLFNNT1YNN L1YGNISASPEEYVA1KAGLHDA1ILRMPHGYD1OVGERGL1SGSGEORVA1IAAI LKDPV1LYDEATSS1DSTTEFT1LGAMDVVRHS1FIARHLS1TVDADETV1DO GKVAERGTHHGL1ANPHS1TSEMH1TQSS1RV0NHD1PKPAKEN1SKBEERK1QOE IYNSVKGCGKSC"
 /db_xref="GI:13623447"
 /db_xref="MIM:603233.1"
 /protein_id="AAH06323.1"
 /cds
 CDS
 BASE COUNT 701 a 465 c 572 g 631 t
 ORIGIN
 Query Match 99.9%; Score 2332; DB 9; Length 2369;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2345; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 QY 1 ATGGCGTGTCTGGGATGATCTTTCGCGCTGGCGCGCGCGCTGCTTTCGAAAG 60

|||||
 Db 11 ATGGCGTGTCTGGGATGATCTTTCGCGCTGGCGCGCGCGCTGCTTTCGAAAG 70
 QY 61 CGCGGACATCCGCGATTCATGATCCGCGCTTATGCTCTGTGACGCGCTCAGGCGCAG 120
 Db 71 CGCGGACATCCGCGATTCATGATCCGCGCTTATGCTCTGTGACGCGCTCAGGCGCAG 130
 QY 121 TGGAGGCAATCACTCGGCGCTTGGAGACCGCTCGAGCTAC--CAGATTCAGAG 177
 Db 131 TGGAGGCAATCACTCGGCGCTTGGAGACCGCTCGAGCTACCGAGATTCAGAG 190
 QY 178 TCATTAAGATATCATATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTATGAT 237
 Db 191 TCATTAAGATATCATATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTATGAT 250
 QY 238 GCGCAAGGCTCTCAGAGATGCGCATGATGAAAGAGCATGTTGGCATGGTCAT 297
 Db 251 GCTGCAAGGCTCTCAGAGATGCGCATGATGAAAGAGCATGTTGGCATGGTCAT 310
 QY 298 GCAGAGAGAGATCCACACAGACCCAAAGAGGTTAAAGATGTTATCTCGAAA 357
 Db 311 GCAGAGAGAGATCCACACAGACCCAAAGAGGTTAAAGATGTTATCTCGAAA 370
 QY 358 ATCATTAAGCAATGCTTTCTTATGTGTGCGCCAAAGACAGCCAGATCTACAGCTAGA 417
 Db 371 ATCATTAAGCAATGCTTTCTTATGTGTGCGCCAAAGACAGCCAGATCTACAGCTAGA 430
 QY 418 GTTGCAATTCGCGGATTTTGGGTGGGCAAGGCCATGAAATTTGGTTCCTTC 477
 Db 431 GTTGCAATTCGCGGATTTTGGGTGGGCAAGGCCATGAAATTTGGTTCCTTC 490
 QY 478 ATGTTTAATATCTGTAGACAGCTCAACAGATGCGGAAACATGCTGAACCTGAGT 537
 Db 491 ATGTTTAATATCTGTAGACAGCTCAACAGATGCGGAAACATGCTGAACCTGAGT 550
 QY 538 GATGCAACCAATACAGTTGCAACCATGCGAACAGAGTTCTGATTGGCTATGCTATCA 597
 Db 551 GATGCAACCAATACAGTTGCAACCATGCGAACAGAGTTCTGATTGGCTATGCTATCA 610
 QY 598 AGAGCTGGAGCTCTTTTAAAGAGTTCGAAAGTTCGAAAGTTCGAAAGTTCGAAAGTTCG 657
 Db 611 AGAGCTGGAGCTCTTTTAAAGAGTTCGAAAGTTCGAAAGTTCGAAAGTTCGAAAGTTCG 670
 QY 658 AATTCATCCGAAGATAGCAAAATGCTTTCTCCATCTTCACAACTGATCTGGGT 717
 Db 671 AATTCATCCGAAGATAGCAAAATGCTTTCTCCATCTTCACAACTGATCTGGGT 730
 QY 718 TTTACCTGAGACAGACAGCGGAGCTTTATCTAAGCTTATGACAGAGAAAGGGGT 777
 Db 731 TTTACCTGAGACAGACAGCGGAGCTTTATCTAAGCTTATGACAGAGAAAGGGGT 790
 QY 778 ATGAGTTTGTCTGAGTGTCTTGTATTAATCTCTCCCATCAATGTTTGAAGTATG 837
 Db 791 ATGAGTTTGTCTGAGTGTCTTGTATTAATCTCTCCCATCAATGTTTGAAGTATG 850
 QY 838 CTGTGATGAGTGTCTTGTATTAATCAATGCGGCGCCAGTTCCTTGGTAACTCTTGA 897
 Db 851 CTGTGATGAGTGTCTTGTATTAATCAATGCGGCGCCAGTTCCTTGGTAACTCTTGA 910
 QY 898 ACACTTGTATCATACAGAGATTACAGTTGACAGTCAACGCTGGAGAACTGATTTAGA 957
 Db 911 ACACTTGTATCATACAGAGATTACAGTTGACAGTCAACGCTGGAGAACTGATTTAGA 970
 QY 958 ATGGAATGAACCAAGACATATGATGACAGTAAATGCTGCTATAGACTACGCTGAAT 1017
 Db 971 ATGGAATGAACCAAGACATATGATGACAGTAAATGCTGCTATAGACTACGCTGAAT 1030
 QY 1018 TATGAACGTGAAGATTTTAAATGAAGATGAGACAGAGATTTGATGATTT 1077
 Db 1031 TATGAACGTGAAGATTTTAAATGAAGATGAGACAGAGATTTGATGATTT 1090
 QY 1078 TTGAAGCGTATGAGCTGCTTCAATGAAAGTACCTTACTGTGCTATGCTGAACCTTT 1137
 |||||

Db	1091	TTGAGACGTATGAGACTGCTTCATTGAAAGTAACCTCTACTCTGCTATGCTGAACCTTT	1150
Qy	1138	GGTCAAGTGCATTTTTCAGTGTGGTTTAAACGCTATATATGTCGTGGCCAGTGAAGGA	1197
Db	1151	GCTCAAGTGCATTTTTCAGTGTGGTTTAAACGCTATATATGTCGTGGCCAGTGAAGGA	1210
Qy	1198	ATTGTGGACGAGTACCTTACTGTTGGAGATCTAGTAAATGTTGAATGAGACTGCTTTTTCAG	1257
Db	1211	ATTGTGGACGAGTACCTTACTGTTGGAGATCTAGTAAATGTTGAATGAGACTGCTTTTTCAG	1270
Qy	1258	CTTTCATTATACCTTCGAACTTTCTGGGAACTGTATATAGAGACTAGACACACTCATTA	1317
Db	1271	CTTTCATTATACCTTCGAACTTTCTGGGAACTGTATATAGAGACTAGACACACTCATTA	1330
Qy	1318	GATATGAACACCTTGTTTACTCTACTCAAGGTAGACACCCAAATTAAGAGACAAATGATG	1377
Db	1331	GATATGAACACCTTGTTTACTCTACTCAAGGTAGACACCCAAATTAAGAGACAAATGATG	1390
Qy	1378	GCATCTCCCTTCAGATCACACACAGACGTAACCGCTTGGATTAATGTGATTTT	1437
Db	1391	GCATCTCCCTTCAGATCACACACAGACGTAACCGCTTGGATTAATGTGATTTT	1450
Qy	1438	GAATACATTGAGGCGCAAGAACTCTTATGTAATATCTTTGAATGCTTGCAGGAAAG	1497
Db	1451	GAATACATTGAGGCGCAAGAACTCTTATGTAATATCTTTGAATGCTTGCAGGAAAG	1510
Qy	1498	AAAGTGGCATTGTAGAGGTAGTGGGTGAGGAAAGCAATGTGAGGCTATATTT	1557
Db	1511	AAAGTGGCATTGTAGAGGTAGTGGGTGAGGAAAGCAATGTGAGGCTATATTT	1570
Qy	1558	CGCTTCTATGAGCTCAAAAAGGGTAGCAATTAATCTTGTGCTCAAAATATACAGATG	1617
Db	1571	CGCTTCTATGAGCTCAAAAAGGGTAGCAATTAATCTTGTGCTCAAAATATACAGATG	1630
Qy	1618	AGCCTGGAAGCTTGGAGGAGTGGAGTGGTACTCAGAGTGTCTCTTCCAT	1677
Db	1631	AGCCTGGAAGCTTGGAGGAGTGGAGTGGTACTCAGAGTGTCTCTTCCAT	1690
Qy	1678	AATATATTTTATTAACACCTCTTATATGGAACATCAGTGTCTCACTGAGGAAAGTAT	1737
Db	1691	AATATATTTTATTAACACCTCTTATATGGAACATCAGTGTCTCACTGAGGAAAGTAT	1750
Qy	1738	GCAGTGGCAAAATTAAGCTGAGCTTATGATGCAATCTTTCGAATGCCATGATATGAC	1797
Db	1751	GCAGTGGCAAAATTAAGCTGAGCTTATGATGCAATCTTTCGAATGCCATGATATGAC	1810
Qy	1798	ACCCAAGTAGGGGAAAGAGAGCTCAAGCTTTCAGAGAGAGAAAGCAAGTAGCAAT	1857
Db	1811	ACCCAAGTAGGGGAAAGAGAGCTCAAGCTTTCAGAGAGAGAAAGCAAGTAGCAAT	1870
Qy	1858	GCAAGAGCATTTTGAAGAGACCCCGCAGTCACTCTATGATGAAGCTTACTTCCTTTA	1917
Db	1871	GCAAGAGCATTTTGAAGAGACCCCGCAGTCACTCTATGATGAAGCTTACTTCCTTTA	1930
Qy	1918	GATTGATTTACTGAAGAGACTTCTTGGTGCATGAAGAGTGTGTCAACACAGAACT	1977
Db	1931	GATTGATTTACTGAAGAGACTTCTTGGTGCATGAAGAGTGTGTCAACACAGAACT	1990
Qy	1978	TCTATTTTCATTGACACACAGATGTCACAGTGTGTTGATGAGATGAATCTTGTCTTG	2037
Db	1991	TCTATTTTCATTGACACACAGATGTCACAGTGTGTTGATGAGATGAATCTTGTCTTG	2050
Qy	2038	GATCAGGGGTAAGGTAGCGGAAGCTGTGTAACCAACATGTTGCTGTAACCTCATGT	2097
Db	2051	GATCAGGGGTAAGGTAGCGGAAGCTGTGTAACCAACATGTTGCTGTAACCTCATGT	2110
Qy	2098	ATCTATTGAGAAATGTGGCATACAGAGACGCTGTGACAGAACATGTAACCCCAA	2157
Db	2111	ATCTATTGAGAAATGTGGCATACAGAGACGCTGTGACAGAACATGTAACCCCAA	2170
Qy	2158	TGGGAAGCAAGAAAGAAATATATCTCAAAAGAGAGAGAAAGAAAGAACTACAGAGAA	2217
Db	2171	TGGGAAGCAAGAAAGAAATATATCTCAAAAGAGAGAGAGAAAGAAAGAACTACAGAGAA	2230

Qy	2218	ATTGTCAATAGTGTGAAAGGCTGTGAAAGCTGTGCTGCTAAGTACATATAGACATTTTC	2277
Db	2231	ATTGTCAATAGTGTGAAAGGCTGTGAAAGCTGTGCTGCTAAGTACATATAGACATTTTC	2290
Qy	2278	TTTTTTTGTGTTTGGACATACATATTTTGCATGACACAGAAATGTTTATTAATAAAAAAT	2337
Db	2291	TTTTTTTGTGTTTGGACATACATATTTTGCATGACACAGAAATGTTTATTAATAAAAAAT	2350
Qy	2338	CATACATT 2345	
Db	2351	CATACATT 2358	
RESULT 4			
AF078777		2333 bp	mRNA
LOCUS			
DEFINITION	Homo sapiens ABC transporter (ATM1) mRNA, partial cds; nuclear gene		
ACCESSION	AF078777		
VERSION	AF078777.1 GI:5702097		
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2333)		
TITLE	Identification of a human mitochondrial ABC transporter, the functional orthologue of yeast Atm1p		
JOURNAL	FEBS Lett. 441 (2), 266-270 (1998)		
MEDLINE	99098366		
PUBMED	9883897		
REFERENCE	2 (bases 1 to 2333)		
AUTHORS	Ceire,P., Hill,R. and Kispal,G.		
TITLE	Direct Submmission		
JOURNAL	Submitted (16-JUL-1998) Cell Biology, University of Marburg, Robert Koch 5, Marburg 35033, Germany		
FEATURES	Location/Qualifiers		
source	1..2333		
gene	/organism="Homo sapiens"		
CDs	/db_xref="taxon:9606"		
	<1..2333		
	/gene="ATM1"		
	<1..2247		
	/gene="ATM1"		
	/note="ATM1p"		
	/product="ABC transporter"		
	/protein_id="A047141.1"		
	/db_xref="GI:5702098"		
	/translation="AMHSRMMAAAAFKRRSAILIRPVSVSGSPQWRPHQLGLA LGTARVQIPESLKSITWRLGNSGQFLDAKAQWFLIKRTCMHAGGLHT DPREGKLDVTRKTIKAMLSYVVKORPDILRAVAISLFLGAKAMNIVPMPFKXA VDSLNMGSMMLISDAENVAATMAAVLIGXSVSAGAAFFNEVRVAIVGKAONS RRLAKVFEHLNLDGPHLSRGTALSKAIDGTGCI SFVLSALVPIMEVWL VSGVLVYKGAQFALVTGLTGTYTAFTVAVTRMRFRIRIEMKADNDAGNAIDSL NRTVATYFNNRREAEORVYDGFELTYTATSLKSTSLAMNPGSALFSVGLTAIMVLA SQGI VAGTLTVGLVWNGLLFOLSLPLNLFGLVYRETRQALIDMNTLPTLLKYDTOI KDRVMA SPLQITPOTATVAVADNVHFEYIEGQKVLSGISFEVPAGKVAIVGSGSGKS TIVRLIFRPEPOKGSYLAGQNIQDVSLSLRAVGVPOQAVLFNITLYNLLXGN ISASPEEVYAVAKLAGLHDAILRMPHGYDVOVERGLKSGGKORVAIARAILKQNP VILIYDEATSLDSTETTELIGAMKDVVKHRTSIFIAHRLSTVADADEIIVLDQKVAE RGTNHLGLANPHSIYSEMHITOSRQNDNPRMEAKENISKEERKKLQDEIIVNSV KGGGNCSC"		
BASE COUNT	688 a	456 c	560 g 629 t
ORIGIN			
Query Match	99.3%	Score 2328.2;	DB 9; Length 2333;
Best Local Similarity	99.98;	Pred. No. 0;	
Matches 2330; Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
Qy	13	GCATGCAATTTTGGCGGTGGCGGCGGCGGCTCTTTCGAAAAGCCGCGCACTCC	72

|||||
Db 1 GCGATGCAATTTGCGCTGGCGCGCGCGGCTGCTTTGCAAAAGCGCGCACCTCC 60
Qy 73 GCGATTCGATCCGGCCCTTTAGTCTCTGTGTAGCGGCTCAGGTCGCGAGTGGAGGCCAAT 132
Db 61 GCGATTCGATCCGGCCCTTTAGTCTCTGTGTAGCGGCTCAGGTCGCGAGTGGAGGCCAAT 120
Qy 133 CAATCGGCGGCTTTGGAAACCGCTCGAGCTTACGAGATTCCAGAGTCAATTAAGATATC 192
Db 121 CAATCGGCGGCTTTGGAAACCGCTCGAGCTTACGAGATTCCAGAGTCAATTAAGATATC 180
Qy 193 ACATGGCAGAGATTGGGAAAAAGGCAATTCAGAGCAATTTAGATGCTGCAAGGCTCTC 252
Db 181 ACATGGCAGAGATTGGGAAAAAGGCAATTCAGAGCAATTTAGATGCTGCAAGGCTCTC 240
Qy 253 CAGGATGGCCCATGATGAGAAAGAGCATGTTGGCATGGTCAATGACGAGAGGAGATC 312
Db 241 CAGGATGGCCCATGATGAGAAAGAGCATGTTGGCATGGTCAATGACGAGAGGAGATC 300
Qy 313 CACACAGACCCCAAGAAAGAGGTTAAAGATGTGATACTCGGAAATCATATAAGCAATG 372
Db 301 CACACAGACCCCAAGAAAGAGGTTAAAGATGTGATACTCGGAAATCATATAAGCAATG 360
Qy 373 CTTTCTTATGTGTGCGCCCAAGACAGGCGCAGATCTAGAGCTTAGAGTTGCCATTTCCGCTG 432
Db 361 CTTTCTTATGTGTGCGCCCAAGACAGGCGCAGATCTAGAGCTTAGAGTTGCCATTTCCGCTG 420
Qy 433 GGATTTTGGGTGGTGCAAAAGGCGATGAATTTGGGTGCTCCCTCAATGTTAAATATGCT 492
Db 421 GGATTTTGGGTGGTGCAAAAGGCGATGAATTTGGGTGCTCCCTCAATGTTAAATATGCT 480
Qy 493 GTAACAGGCTTCAACAGATGTTCGGGAAACATGCTGAACCTGATGATGACCAAAATACA 552
Db 481 GTAACAGGCTTCAACAGATGTTCGGGAAACATGCTGAACCTGATGATGACCAAAATACA 540
Qy 553 GTTGCAACCATGGCAACAGCAGTTCTGATTTGGCTATGGTGTATCAAGAGCTGGAGCTGCT 612
Db 541 GTTGCAACCATGGCAACAGCAGTTCTGATTTGGCTATGGTGTATCAAGAGCTGGAGCTGCT 600
Qy 613 TTTTAAACGAAGTTGCAAAATGCAATTTGGCAAGGAGCCGGAATTCATCCGANA 672
Db 601 TTTTAAACGAAGTTGCAAAATGCAATTTGGCAAGGAGCCGGAATTCATCCGANA 660
Qy 673 ATAGCCAAAAATGTCTTCTTCATCTTCAACAACCTGATCTGGGTTTTCACCTAGAGCA 732
Db 661 ATAGCCAAAAATGTCTTCTTCATCTTCAACAACCTGATCTGGGTTTTCACCTAGAGCA 720
Qy 733 CAGACGGAGCTTTATCTAAGGCTATTGACAGAGAAACAGGGGTATCAGTTTGTCTG 792
Db 721 CAGACGGAGCTTTATCTAAGGCTATTGACAGAGAAACAGGGGTATCAGTTTGTCTG 780
Qy 793 AGTGCTTTGGATTTAATCTTCTCCCATCATGTTTGAAGGATGCTTGTCAAGTGTGTT 852
Db 781 AGTGCTTTGGATTTAATCTTCTCCCATCATGTTTGAAGGATGCTTGTCAAGTGTGTT 840
Qy 853 TTGTATTACAAATCGGGTGCCAGTTTGTGTTAACCCTTGGAAACATTTGATACATAC 912
Db 841 TTGTATTACAAATCGGGTGCCAGTTTGTGTTAACCCTTGGAAACATTTGATACATAC 900
Qy 913 ACAGCATTCACAGTTGACATCAACGCTGAGAACTAGATTAGAAATGAAACAAA 972
Db 901 ACAGCATTCACAGTTGACATCAACGCTGAGAACTAGATTAGAAATGAAACAAA 960
Qy 973 GCGATATATGATGAGGTAATGCTGCTATAGACTACTGCTGAATTTAGAACTGTGAAG 1032
Db 961 GCGATATATGATGAGGTAATGCTGCTATAGACTACTGCTGAATTTAGAACTGTGAAG 1020
Qy 1033 TATTTATATATGAAGATATGAAGACAGAGATATGATGATTTTGAAGACGTATGAG 1092
Db 1021 TATTTATATATGAAGATATGAAGACAGAGATATGATGATTTTGAAGACGTATGAG 1080
Qy 1093 ACTGCTTCAATGAAAGTACCTTACTCTGGCTATGCTGAACTTTGGTCAAAAGTCTATT 1152
|||||

Db 1081 ACTGCTTCAATGAAAGTACCTTACTCTGGCTATGCTGAACTTTGGTCAAAAGTCTATT 1140
Qy 1153 TTGAGTTCGGGTTTAAACAGCTATATATGCTGCTGCCAGTACAGGAAATTTGGCAGTACC 1212
Db 1141 TTGAGTTCGGGTTTAAACAGCTATATATGCTGCTGCCAGTACAGGAAATTTGGCAGTACC 1200
Qy 1213 CTTACTGTGAGATCTAGTATATGTTGAATGACCTGCTTTTTCAGCTTTCAATTAACCCCTG 1272
Db 1201 CTTACTGTGAGATCTAGTATATGTTGAATGACCTGCTTTTTCAGCTTTCAATTAACCCCTG 1260
Qy 1273 AACTTTTGGGAACTGTATATAGAGACCTAGCAAGCACTCATAGTATAGAACCTTTG 1332
Db 1261 AACTTTTGGGAACTGTATATAGAGACCTAGCAAGCACTCATAGTATAGAACCTTTG 1320
Qy 1333 TTTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGGTGATGCTCCCTCAG 1392
Db 1321 TTTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGGTGATGCTCCCTCAG 1380
Qy 1393 ATCACACACAGACAGCTACCGTGCCCTTGTATATGTGATTTTGAATACATTTAGAGGC 1452
Db 1381 ATCACACACAGACAGCTACCGTGCCCTTGTATATGTGATTTTGAATACATTTAGAGGC 1440
Qy 1453 CAGAAAGTCTTATGTTGAATATCTTTGAAGTCCCTGACAGAAAGAAAGTGGCCATTGTA 1512
Db 1441 CAGAAAGTCTTATGTTGAATATCTTTGAAGTCCCTGACAGAAAGAAAGTGGCCATTGTA 1500
Qy 1513 GGAAGTATGAGGTACAGGAAAGACAAATATGAGAGCTATATTTGGCTCTATGAGCCT 1572
Db 1501 GGAAGTATGAGGTACAGGAAAGACAAATATGAGAGCTATATTTTGGCTCTATGAGCCT 1560
Qy 1573 CAAAAGGTTAGCAATTAATCTTGTGTCCTCAAAATATACAAAGATGAGCCTGAAAGCCTT 1632
Db 1561 CAAAAGGTTAGCAATTAATCTTGTGTCCTCAAAATATACAAAGATGAGCCTGTAAGCCTT 1620
Qy 1633 CGAGAGGCAATGGGAGTGTATCTCAGAGATGCTGTCTCTTCCATATATCTATTTATTAC 1692
Db 1621 CGAGAGGCAATGGGAGTGTATCTCAGAGATGCTGTCTCTTCCATATATCTATTTATTAC 1680
Qy 1693 AACCTTTATATGGAACATCAGTGGCTTCACTGAGGAAGTGTATGAGTGGGCAAAATTA 1752
Db 1681 AACCTTTATATGGAACATCAGTGGCTTCACTGAGGAAGTGTATGAGTGGGCAAAATTA 1740
Qy 1753 GCTGCACTTCATGATGCAATTTCTTGCATATGCCATATGATATGACCCAAATGAGGAA 1812
Db 1741 GCTGCACTTCATGATGCAATTTCTTGCATATGCCATATGATATGACCCAAATGAGGAA 1800
Qy 1813 CGAGACTCAAGCTTTTCAGAGAGGAAAGCAAAAGATGACCAATTTGACAGCCATTTTG 1872
Db 1801 CGAGACTCAAGCTTTTCAGAGAGGAAAGCAAAAGATGACCAATTTGACAGCCATTTTG 1860
Qy 1873 AAGGACCCCGCAATCTATCTATATGATGAGCTAATCTTCAATCGTTAATGATTACTGAA 1932
Db 1861 AAGGACCCCGCAATCTATCTATATGATGAGCTAATCTTCAATCGTTAATGATTACTGAA 1920
Qy 1933 GAGACTATCTTGGTGCATGAAGATGTGTCAAAACACAGAACTTCTATTTCAATTTGCA 1992
Db 1921 GAGACTATCTTGGTGCATGAAGATGTGTCAAAACACAGAACTTCTATTTCAATTTGCA 1980
Qy 1993 CACAGATTGTCAACAGTGTGATGACAGATGAATCATTTGTTGATCAAGGTTAAGTGA 2052
Db 1981 CACAGATTGTCAACAGTGTGATGACAGATGAATCATTTGTTGATCAAGGTTAAGTGA 2040
Qy 2053 GCGGACGTGATGCCACCAATGGTTTGTGCTTAACCTCATAGTATCTATTCAGAAATG 2112
Db 2041 GCGGACGTGATGCCACCAATGGTTTGTGCTTAACCTCATAGTATCTATTCAGAAATG 2100
Qy 2113 TGGCATATCACAGAGACCGGTGTGAGAAACATGATTAACCCCAATGGGAAAGCAAGAAA 2172
Db 2101 TGGCATATCACAGAGACCGGTGTGAGAAACATGATTAACCCCAATGGGAAAGCAAGAAA 2160
Qy 2173 GAAATATATCCAAAGAGAGGAAAGAAAGAACTACAAAGAAATTTGCTCAATAGTGTG 2232
Db 2161 GAAATATATCCAAAGAGAGGAAAGAAAGAACTACAAAGAAATTTGCTCAATAGTGTG 2220
|||||

2223 AAAGCTGTGGAAGCTTTGCTAGTACATAGACATTTCTTTTGTGTTT 2292
Db 2221 AAAGCTGTGGAAGCTTTGCTAGTACATAGACATTTCTTTTGTGTTT 2280
Qy 2293 GGACTACATATTTGACATGAGCAGATTTGTTTATTTAAAAATCATATTT 2345
Db 2281 GGACTACATATTTGACATGAGCAGATTTGTTTATTTAAAAATCATATTT 2333

RESULT 5
AR028561 2407 bp DNA linear PAT 29-SEP-1999
LOCUS AR028561
DEFINITION Sequence 2 from patent US 5858719.
ACCESSION AR028561
VERSION AR028561.1 GI:5940534
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2407)
AUTHORS Hillman, J.L., Shah, P. and Corley, N.C.
TITLE Polynucleotides encoding human ATP binding-casette transport
JOURNAL Patent: US 5858719-A 2 12-JAN-1999;
FEATURES
source location/Qualifiers
1..2407
/organism="unknown"
BASE COUNT 726 a 477 c 570 g 629 t 5 others
ORIGIN

Query Match 95.9%; Score 2248; DB 6; Length 2407;
Best local similarity 99.4%; Pred. No. 0; Mismatches 13; Indels 1; Gaps 1;
Matches 2265; Conservative 0; Nucleotides 13; Indels 1; Gaps 1;

Qy 8 TCGTCGCATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAACGCCGCGC 67
Db 1 TCGTCGCATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAACGCCGCGC 60
Qy 68 ACTCCGCGATTTGATTCGCGCTTTAGTCTCTGTTAGCGGCTCAGTCCGCGTGAAGC 127
Db 61 ACTCCGCGATTTGATTCGCGCTTTAGTCTCTGTTAGCGGCTCAGTCCGCGTGAAGC 120
Qy 128 CACATCAACTCGGCGCTTGGGAAACCGCTCGAGCTTCCAGATTCGAGATCATTAATAA 187
Db 121 CACATCAACTCGGCGCTTGGGAAACCGCTCGAGCTTCCAGATTCGAGATCATTAATAA 180
Qy 188 GTATCAATGCGAGATTTGGGAAAAGGCAATTCAGACAGTTCCTAGATGCTGCAAGG 247
Db 181 GTATCAATGCGAGATTTGGGAAAAGGCAATTCAGACAGTTCCTAGATGCTGCAAGG 240
Qy 248 CTCTCAGGTATGCGCATGATGAAAAGAGACATGTTGGCATGCTCAGAGAGAG 307
Db 241 CTCTCAGGTATGCGCATGATGAAAAGAGACATGTTGGCATGCTCAGAGAGAG 300
Qy 308 GACTCCACACAGACCCAAAGAGGTTAAAGATGTTGATCTCGGAAAATCATTAAG 367
Db 301 GACTCCACACAGACCCAAAGAGGTTAAAGATGTTGATCTCGGAAAATCATTAAG 360
Qy 368 CAATCTTTCTATGTTGGCCCAAGACAGGCTCAGATTCAGAGCTAGAGTTGCATTT 427
Db 361 CAATCTTTCTATGTTGGCCCAAGACAGGCTCAGATTCAGAGCTAGAGTTGCATTT 420
Qy 428 CGCTGGGATTTTGGGTGTGCAAAAGCCATATATTTGGTCTCTTCATGTTTAAT 487
Db 421 CGCTGGGATTTTGGGTGTGCAAAAGCCATATATTTGGTCTCTTCATGTTTAAT 480
Qy 488 ATGCTGTGAGAGCTTCAACGAGATGCGGAAAACATGCTGAACCTGAGTATGACCAA 547
Db 481 ATGCTGTGAGAGCTTCAACGAGATGCGGAAAACATGCTGAACCTGAGTATGACCAA 540
Qy 548 ATACAGTTGCAACATGACAGAGCTTCTGATTTGCTATGCTATCAAGCTGAG 607

Db 541 ATACAGTTGCAACCATGCGCAACAGAGTTCGATTTGCTATGCTATCAAGCTGAG 600
Qy 608 CTGCTTTTAAAGAGTTGGAATGCAATGTTGGCAAGTATGCCAGAAATCAATCC 667
Db 601 CTGCTTTTAAAGAGTTGGAATGCAATGTTGGCAAGTATGCCAGAAATCAATCC 660
Qy 668 GAAGAATAGCCAAAATGCTTTCTCCATCTTCAACACCTGATCTGGTTTCACTGA 727
Db 661 GAAGAATAGCCAAAATGCTTTCTCCATCTTCAACACCTGATCTGGTTTCACTGA 720
Qy 728 GCAGACAGACGGAGCTTATCTAAGGCTATGACAGAGAAACAGGGTATCGTTTG 787
Db 721 GCAGACAGACGGAGCTTATCTAAGGCTATGACAGAGAAACAGGGTATCGTTTG 780
Qy 788 TCTGAGTCTTGGTATTTAATCTCTCCATCATGTTGAAGTATGATGCTGCTGAGT 847
Db 781 TCTGAGTCTTGGTATTTAATCTCTCCATCATGTTGAAGTATGATGCTGCTGAGT 840
Qy 848 GTGTTTGTATTAAGAAATGCGGTCGCCAGTTTGTGTTGTAACCTTGAAACATTTGTA 907
Db 841 GTGTTTGTATTAAGAAATGCGGTCGCCAGTTTGTGTTGTAACCTTGAAACATTTGTA 900
Qy 908 CATACACAGCATTCACAGTTGCACTACACGCTGAGAACTAGATTGAATGAAATGA 967
Db 901 CATACACAGCATTCACAGTTGCACTACACGCTGAGAACTAGATTGAATGAAATGA 960
Qy 968 ACAAGCAGATTAATGATGCGGTAATGCTGCTATGACTACGCTGTAATGAAATG 1027
Db 961 ACAAGCAGATTAATGATGCGGTAATGCTGCTATGACTACGCTGTAATGAAATG 1020
Qy 1028 TGAAGTATTTAATGAATGAATATGAAACAGACAGATATGATGATTTTGAAGAGCT 1087
Db 1021 TGAAGTATTTAATGAATGAATATGAAACAGACAGATATGATGATTTTGAAGAGCT 1080
Qy 1088 ATGAGACTGCTTCAATGAAAAGTACTCTGCTATGCTGCTGCTGCTGCTGCTGCTG 1147
Db 1081 ATGAGACTGCTTCAATGAAAAGTACTCTGCTATGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1148 CTATTTTCAAGTGTGTTTAAACAGTAAATGAGTGTGCTGCGAGTCAAGGAAATGTCGAG 1207
Db 1141 CTATTTTCAAGTGTGTTTAAACAGTAAATGAGTGTGCTGCGAGTCAAGGAAATGTCGAG 1200
Qy 1208 GTACCTTACTGTTGAGATCTAGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1267
Db 1201 GTACCTTACTGTTGAGATCTAGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1260
Qy 1268 CCTGAACTTTCTGAGAACTGTATATAGAGACTAGACAAAGCTCATATATATGAACA 1327
Db 1261 CCTGAACTTTCTGAGAACTGTATATAGAGACTAGACAAAGCTCATATATATGAACA 1320
Qy 1328 CTTGTTTAACTCTAAGTATGAGTAAACCCAAATTTAAAGCAAGTATGTCATCTCC 1387
Db 1321 CTTGTTTAACTCTAAGTATGAGTAAACCCAAATTTAAAGCAAGTATGTCATCTCC 1380
Qy 1388 TTCAGATCACACACAGACAGTACCGTGGCTTGTGAATGTCATTTGAATACATTTG 1447
Db 1381 TTCAGATCACACACAGACAGTACCGTGGCTTGTGAATGTCATTTGAATACATTTG 1440
Qy 1448 AGGGCAGAAAGTCTTATAGTAAATCTTTGAAGTCTGACAGAAAGAAAGTGGCCA 1507
Db 1441 AGGGCAGAAAGTCTTATAGTAAATCTTTGAAGTCTGACAGAAAGAAAGTGGCCA 1500
Qy 1508 TTGTAGAGAGTATGAGGTCAAGGAAAAGCAATATAGTGAAGCTATTTATTTGCTTATG 1567
Db 1501 TTGTAGAGAGTATGAGGTCAAGGAAAAGCAATATAGTGAAGCTATTTATTTGCTTATG 1560
Qy 1568 AGCCTCAAAAGGTAGCATTTATCTGCTGCTCAAAATATCAAGATGTCAGCTGAAA 1627
Db 1561 AGCCTCAAAAGGTAGCATTTATCTGCTGCTCAAAATATCAAGATGTCAGCTGAAA 1620
Qy 1628 GCCTTGGAGGGCAGTGGAGTGTACTCAGATGCTGTCCTTCCATATATCTATTT 1687
Db 1621 GCCTTGGAGGGCAGTGGAGTGTACTCAGATGCTGTCCTTCCATATATCTATTT 1680

QY 1688 ATTCACACCTCTTATATGAGAAATCACTAGCTTCACTGAGAAATGTATGCACTGGCAA 1747
DB 1681 ATTACAACTCTTATATGAGAAATCACTAGCTTCACTGAGAAATGTATGCACTGGCAA 1740
QY 1748 AATTAGCTGACCTTATATGAGAAATCACTAGCTTCACTGAGAAATGTATGCACTGGCAA 1807
DB 1741 AATTAGCTGACCTTATATGAGAAATCACTAGCTTCACTGAGAAATGTATGCACTGGCAA 1800
QY 1808 GGAACGAGACTCAAGCTTCAAGAGAGAAAGCAAGAGTGTGCAATTCAGAGGCA 1867
DB 1801 GGAACGAGACTCAAGCTTCAAGAGAGAAAGCAAGAGTGTGCAATTCAGAGGCA 1860
QY 1868 TTTTGAAGACCCCGCACTACTCTATGATGAAAGCTTCACTGATTCGTTAGATTGATTA 1927
DB 1861 TTTTGAAGACCCCGCACTACTCTATGATGAAAGCTTCACTGATTCGTTAGATTGATTA 1920
QY 1928 CTGAAGAGACTTCTTGTGTCCTGATGAAAGAGTGTGCAACACAGAACTTCTATTTTCA 1987
DB 1921 CTGAAGAGACTTCTTGTGTCCTGATGAAAGAGTGTGCAACACAGAACTTCTATTTTCA 1980
QY 1988 TTGCACACAGATTGTGCAACAGTGTGATGAGATGAAATCATTTGTGTGATGAGGTA 2047
DB 1981 TTGCACACAGATTGTGCAACAGTGTGATGAGATGAAATCATTTGTGTGATGAGGTA 2040
QY 2048 AGGTAGCCGAACTGTGTAACCCACCATGTTGCTTGTCAACCTCATAGTATCTATTCA 2107
DB 2041 AGGTAGCCGAACTGTGTAACCCACCATGTTGCTTGTCAACCTCATAGTATCTATTCA 2100
QY 2108 AATATGTCATACACAGAGCCGCTGTGCAAACTATGATTAACCCCAATGGAAGCAA 2167
DB 2101 AATATGTCATACACAGAGCCGCTGTGCAAACTATGATTAACCCCAATGGAAGCAA 2160
QY 2168 AGAAGAAATATATCCAAAGAGAGAAAGAAAGAACTCAAGAGAAATTCATATA 2227
DB 2161 AGAAGAAATATATCCAAAGAGAGAAAGAAAGAACTCAAGAGAAATTCATATA 2220
QY 2228 GTGTGAAGAGCTGTGAAATGTTCTGTCTAGTCACTAAGACATTTTCTTTTGT 2286
DB 2221 GTGTGAAGAGCTGTGAAATGTTCTGTCTAGTCACTAAGACATTTTCTTTTGT 2278

RESULT 6
AR100662 2407 bp DNA linear PAT 14-FEB-2001
LOCUS AR100662
DEFINITION Sequence 2 from patent US 6080842.
ACCESSION AR100662
VERSION AR100662.1 GI:12811110
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2407)
AUTHORS Hillman,J.L., Shah,P. and Corley,N.C.
TITLE Human ATP binding-cassette transport protein
JOURNAL Patent: US 6080842-A 2 27-JUN-2000;
FEATURES
source location/Qualifiers
1..2407
/organism="unknown"
BASE COUNT 726 a 477 c 570 g 629 t 5 others
ORIGIN

Query Match 95.9%; Score 2248; DB 6; Length 2407;
Best local Similarity 99.4%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 8 TGCTCGGATGATCTTGTGGCGTGGCGGCGCGCGCTGCTTTGAAAGCGCCGCG 67
DB 1 TGCTCGGATGATCTTGTGGCGTGGCGGCGCGCGCTGCTTTGAAAGCGCCGCG 60
QY 68 ACTCGGATGATCTTGTGGCGTGGCGGCGCGCTGCTTTGAAAGCGCCGCG 127
DB 61 ACTCGGATGATCTTGTGGCGTGGCGGCGCGCTGCTTTGAAAGCGCCGCG 120

QY 128 CACATCAACTCGGCGCTTGGGAACCGCTTCAGCTTACAGATTCAGAGCTATTAATAA 187
DB 121 CACATCAACTCGGCGCTTGGGAACCGCTTCAGCTTACAGATTCAGAGCTATTAATAA 180
QY 188 GTATCATGAGGACAGATTTGGGAAAGGCAATTCAGAGAGCTTTATGATGCTGCAAG 247
DB 181 GTATCATGAGGACAGATTTGGGAAAGGCAATTCAGAGAGCTTTATGATGCTGCAAG 240
QY 248 CTCTCCAGGATGAGGCACTGATGAAAGAGGCAATGTTGGCATGCGAGAGAGAG 307
DB 241 CTCTCCAGGATGAGGCACTGATGAAAGAGGCAATGTTGGCATGCGAGAGAGAG 300
QY 308 GACTCCACAGAGCCCAAGAGAGGTTAAAGATGTTGATCTCGGAAATCATTAAG 367
DB 301 GACTCCACAGAGCCCAAGAGAGGTTAAAGATGTTGATCTCGGAAATCATTAAG 360
QY 368 CAATGCTTTCTTATGTGTGCGCCCAAGAGAGGCTGATCTAGAGCTTAGAGTTGCCATTT 427
DB 361 CAATGCTTTCTTATGTGTGCGCCCAAGAGAGGCTGATCTAGAGCTTAGAGTTGCCATTT 420
QY 428 CGCTGGGATTTTGGGTGTGCAAGGCGCATGATATGTGTTCCCTTCATGTTTAAT 487
DB 421 CGCTGGGATTTTGGGTGTGCAAGGCGCATGATATGTGTTCCCTTCATGTTTAAT 480
QY 488 ATGCTGAGACAGCTCAACAGATGTCGGAAGACATGCTGAACCTGATGATGACCAA 547
DB 481 ATGCTGAGACAGCTCAACAGATGTCGGAAGACATGCTGAACCTGATGATGACCAA 540
QY 548 ATACAGTTGCAACCATGCAACAGAGTTCTGATGTTGCTATGTTGATCAAGAGCTGAG 607
DB 541 ATACAGTTGCAACCATGCAACAGAGTTCTGATGTTGCTATGTTGATCAAGAGCTGAG 600
QY 608 CTGCTTTTCTTAAAGAGTTCGAAATGCAATTTGGCAAGGATGACCAATTCATTC 667
DB 601 CTGCTTTTCTTAAAGAGTTCGAAATGCAATTTGGCAAGGATGACCAATTCATTC 660
QY 668 GAAGAAATGCAAAATATGTTCTTCTCATCTTCAACAGCTGATCTGGTTTCACTGA 727
DB 661 GAAGAAATGCAAAATATGTTCTTCTCATCTTCAACAGCTGATCTGGTTTCACTGA 720
QY 728 GCAGACAGAGCGGAGCTTATCTAAGGCTATGACAGAGAACAGGAGTATCAGTTTGT 787
DB 721 GCAGACAGAGCGGAGCTTATCTAAGGCTATGACAGAGAACAGGAGTATCAGTTTGT 780
QY 788 TCTTGAGTGTGTTGATTTTAATCTTCTCCATCATGTTTGAAGATGCTGTCAAG 847
DB 781 TCTTGAGTGTGTTGATTTTAATCTTCTCCATCATGTTTGAAGATGCTGTGTCAAG 840
QY 848 GTGTTTGTATTAATAATGCGGAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 907
DB 841 GTGTTTGTATTAATAATGCGGAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
QY 908 CATACACAGATTCACAGTTGCAAGTCAACAGGTCAGAGAACTAGATTTAGATGAAGATGA 967
DB 901 CATACACAGATTCACAGTTGCAAGTCAACAGGTCAGAGAACTAGATTTAGATGAAGATGA 960
QY 968 ACAAGCAGATTAATGATGAGGTTAATGCTGCTATTAAGTCACTGCTGAATTAAGAACTG 1027
DB 961 ACAAGCAGATTAATGATGAGGTTAATGCTGCTATTAAGTCACTGCTGAATTAAGAACTG 1020
QY 1028 TGAAGTATTTTAATAATGAAAGATGAAAGCAGAGATATGATGATTTTGAAGCGT 1087
DB 1021 TGAAGTATTTTAATAATGAAAGATGAAAGCAGAGATATGATGATTTTGAAGCGT 1080
QY 1088 ATGAGACTGCTTCATTTGAAGAGTACTCTGCTATGCTGAACCTTTGATCAAGATG 1147
DB 1081 ATGAGACTGCTTCATTTGAAGAGTACTCTGCTATGCTGAACCTTTGATCAAGATG 1140
QY 1148 CTATTTTCAAGTGTGCTTTTAAACAGCTATAATGTTGCTGCGCACTCAGGGAATTTGCGCAG 1207
DB 1141 CTATTTTCAAGTGTGCTTTTAAACAGCTATAATGTTGCTGCGCACTCAGGGAATTTGCGCAG 1200

```

QY 1208 GTACCTTAAGTGTGAGATCTAGTAATGTGAATGAGACTGCTTTTTCAGCTTTCAATTAAC 1267
    |||
    |||
    |||
Db 1201 GTACCTTAAGTGTGAGATCTAGTAATGTGAATGAGACTGCTTTTTCAGCTTTCAATTAAC 1260
    |||
    |||
    |||
QY 1268 CCTTGAACCTTTCTGGGAAGTGTATATAGAGACTAGACACTCATAGATATGAGAA 1327
    |||
    |||
    |||
Db 1261 CCTTGAACCTTTCTGGGAAGTGTATATAGAGACTAGACACTCATAGATATGAGAA 1320
    |||
    |||
    |||
QY 1328 CTTTGTACTCTACTCAAGTAGAGACCCCAATTAAGACAAAGTAGATGATCCTCC 1387
    |||
    |||
    |||
Db 1321 CTTTGTACTCTACTCAAGTAGAGACCCCAATTAAGACAAAGTAGATGATCCTCC 1380
    |||
    |||
    |||
QY 1388 TTCAATGATCACACAGACAGCTACCGTGGCTTTGATATGTGATTTGAATATG 1447
    |||
    |||
    |||
Db 1381 TTCAATGATCACACAGACAGCTACCGTGGCTTTGATATGTGATTTGAATATG 1440
    |||
    |||
    |||
QY 1448 AGGGCCAGAAATCTCTAGTGTGAATATCTTTGAAGTCCCTGACAGAAAGAGTGGCCA 1507
    |||
    |||
    |||
Db 1441 AGGGCCAGAAATCTCTAGTGTGAATATCTTTGAAGTCCCTGACAGAAAGAGTGGCCA 1500
    |||
    |||
    |||
QY 1508 TTGTAGAGAGTGTGGGTGAGGAGAAAGACAAATAGTGAAGCTATTTATTTGCTTTATG 1567
    |||
    |||
    |||
Db 1501 TTGTAGAGAGTGTGGGTGAGGAGAAAGACAAATAGTGAAGCTATTTATTTGCTTTATG 1560
    |||
    |||
    |||
QY 1568 AGCTTCAAAAGGGTAGACATTTATCTTGTCTGTCAAAATATACAGATGTGACCTGGAAA 1627
    |||
    |||
    |||
Db 1561 AGCTTCAAAAGGGTAGACATTTATCTTGTCTGTCAAAATATACAGATGTGACCTGGAAA 1620
    |||
    |||
    |||
QY 1628 GCTTTCGAGAGGAGTGTGGAGAGTGTACCTCAGAGATGTCTCTCTTCCATATCTATTT 1687
    |||
    |||
    |||
Db 1621 GCTTTCGAGAGGAGTGTGGAGAGTGTACCTCAGAGATGTCTCTCTTCCATATCTATTT 1680
    |||
    |||
    |||
QY 1688 ATTACAACTCTTATATGAGAAACATCAGTGTCTACCTGAGAGAGTGTGACAGTGAGCA 1747
    |||
    |||
    |||
Db 1681 ATTACAACTCTTATATGAGAAACATCAGTGTCTACCTGAGAGAGTGTGACAGTGAGCA 1740
    |||
    |||
    |||
QY 1748 AATTAGCTGAGCTTCATGATGATGATTTCTTGAATGACACATGATATGACACCAAGTAG 1807
    |||
    |||
    |||
Db 1741 AATTAGCTGAGCTTCATGATGATGATTTCTTGAATGACACATGATATGACACCAAGTAG 1800
    |||
    |||
    |||
QY 1808 GGGAGCAGAGACTCAAGCTTTTCAGAGAGAGAAAGCAAGATGACATTTGCAAGAGCCA 1867
    |||
    |||
    |||
Db 1801 GGGAGCAGAGACTCAAGCTTTTCAGAGAGAGAAAGCAAGATGACATTTGCAAGAGCCA 1860
    |||
    |||
    |||
QY 1868 TTTTGAAGAGCCCCCAGTCAATCTATGATGAGAGCTACTTCATGCTAGATGCAATTA 1927
    |||
    |||
    |||
Db 1861 TTTTGAAGAGCCCCCAGTCAATCTATGATGAGAGCTACTTCATGCTAGATGCAATTA 1920
    |||
    |||
    |||
QY 1928 CTGAAGAGACTATCTTGTGTGCTAGTAAGATGTGTGCAACAGACAACTTCTATTTTCA 1987
    |||
    |||
    |||
Db 1921 CTGAAGAGACTATCTTGTGTGCTAGTAAGATGTGTGCAACAGACAACTTCTATTTTCA 1980
    |||
    |||
    |||
QY 1988 TTGCAACAGATTTGTCAACAGTGTGTGATGCAATGTAATCATTTGTTGGATCAGGGTA 2047
    |||
    |||
    |||
Db 1981 TTGCAACAGATTTGTCAACAGTGTGTGATGCAATGTAATCATTTGTTGGATCAGGGTA 2040
    |||
    |||
    |||
QY 2048 AGTGAAGCCGAAGTGTGATCCCAACATGTTTGTCTGCTAAGCTCAATAGTATCTATTTCA 2107
    |||
    |||
    |||
Db 2041 AGTGAAGCCGAAGTGTGATCCCAACATGTTTGTCTGCTAAGCTCAATAGTATCTATTTCA 2100
    |||
    |||
    |||
QY 2108 AAATGTGCGATACAGAGACAGCCGTGTGACAGAACCATATGTAACCCCAATGGAGAGCA 2167
    |||
    |||
    |||
Db 2101 AAATGTGCGATACAGAGACAGCCGTGTGACAGAACCATATGTAACCCCAATGGAGAGCA 2160
    |||
    |||
    |||
QY 2168 AGAAAGAAATATATTCAGAAAGAGAGAGAAAGAAATTAACAAGAAATTTGTCAATA 2227
    |||
    |||
    |||
Db 2161 AGAAAGAAATATATTCAGAAAGAGAGAGAAAGAAATTAACAAGAAATTTGTCAATA 2220
    |||
    |||
    |||
QY 2228 GTGTGAAGAGCTGTGGAAGCTGTGTGCTAAGTCAATAGACATTTTCTTTTGTGT 2286
    |||
    |||
    |||
Db 2221 GTGTGAAGAGCTGTGGAAGCTGTGTGCT -AGTCAATAGACATTTCTTTTGTGT 2278
    |||
    |||
    |||

```

```

RESULT 7
AK001418      2193 bp      mRNA      linear      PRI 01-AUG-2002
LOCUS         Homo sapiens cDNA FLJ10556 fis, clone NT2RP2002479, highly similar
DEFINITION   to Homo sapiens mRNA for ABC transporter 7 protein.
ACCESSION    AK001418
VERSION      AK001418.1 GI:7022662
KEYWORDS     oligo capping; fis (full insert sequence).
SOURCE       Homo sapiens testocarcinoma cell line:NT2 cDNA to mRNA,
              clone 11b:NT2RP2 clone:NT2RP2002479.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1
AUTHORS      Isegai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
              Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
              Tanase,T., Nomura,Y., Togiyu,S., Komai,F., Hara,R., Takeuchi,K.,
              Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
              Wakamatsu,A., Nakamura,Y., Nagaiari,K., Masuno,Y. and Ohshima,A.
              NEDO human cDNA sequencing project
              Unpublished
              2 (bases 1 to 2193)
TITLE        Isegai,T. and Otsuki,T.
JOURNAL      Direct Submission
AUTHORS      Submitted (16-FEB-2000) Takao Isegai, Helix Research Institute,
              Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
              NEDO human cDNA sequencing project supported by Ministry of
              International Trade and Industry of Japan. cDNA full insert
              sequencing: Research Association for Biotechnology; cDNA library
              construction; 5'- & 3'-end one pass sequencing and clone selection:
              Helix Research Institute (supported by Japan Key Technology Center
              etc.) and Department of Virology, Institute of Medical Science,
              University of Tokyo.
FEATURES     Location/Qualifiers
              source          1..2193
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
                           /clone="NT2RP2002479"
                           /cell_line="NT2"
                           /cell_type="teratocarcinoma"
                           /clone_11b="NT2RP2"
                           /note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal
                           precursor cells after 2-weeks retinoic acid (RA)
                           induction."
BASE COUNT   621 a      451 c      536 g      585 t
ORIGIN
Query Match          91.7%; Score 2150.4; DB 9; Length 2193;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 ATGGGCTGCTCGGATGATCTTTTGGGCGTGGGCGGCGGCGGCTTTGGAAAG 60
    |||
    |||
    |||
Db 26 ATGGGCTGCTCGGATGATCTTTTGGGCGTGGGCGGCGGCGGCGGCTTTGGAAAG 85
    |||
    |||
    |||
QY 61 CGCGGCACTCCGCGATTTCTGATCCGCGCTTATGCTCTGTAGACGGCTCAGTCCGAG 120
    |||
    |||
    |||
Db 86 CGCGGCACTCCGCGATTTCTGATCCGCGCTTATGCTCTGTAGACGGCTCAGTCCGAG 145
    |||
    |||
    |||
QY 121 TGAAGCCACATCAACTGGGCGCTTGGGAACGCTCGAGCTTAC--CAGATTCAGAG 177
    |||
    |||
    |||
Db 146 TGAAGCCACATCAACTGGGCGCTTGGGAACGCTCGAGCTTAC--CAGATTCAGAG 205
    |||
    |||
    |||
QY 178 TCATTAAAAAGTATCACATGGCAGAGATTTGGGAAAGGCAATTTCAGACAGTTCTTAGAT 237
    |||
    |||
    |||
Db 206 TCATTAAAAAGTATCACATGGCAGAGATTTGGGAAAGGCAATTTCAGACAGTTCTTAGAT 265
    |||
    |||
    |||
QY 238 GGTGCAAGAGCTTCGAGGTATGGCACTGATAGAAAAGAGACATGTGTGATGATGAT 297
    |||
    |||
    |||
Db 266 GGTGCAAGAGCTTCGAGGTATGGCACTGATAGAAAAGAGACATGTGTGATGATGAT 325
    |||
    |||
    |||
QY 298 GCAGAGAGAGACTTCCACACAGACCCAAAGAGGTTTAAAGATGTTGATATCTCGAAA 357
    |||
    |||
    |||

```

Db 326 GCAGAGAGAGAGCTCCACAGAGCCCAAAAGGGTTAAAAGTGTGATATCGGAAA 385
 Qy 358 ATCTAAAAAGCATGCTTTCTTATGTGTGGCCCAAAAGCAGCCAGATCTACGAGCTAGA 417
 Db 386 ATCTAAAAAGCATGCTTTCTTATGTGTGGCCCAAAAGCAGCCAGATCTACGAGCTAGA 445
 Qy 418 GTTCCCATTTGCTGGGATTTTGGGTGGTCCAAAGCCCATGAATATTTGGTTCCTTC 477
 Db 446 GTTCCCATTTGCTGGGATTTTGGGTGGTCCAAAGCCCATGAATATTTGGTTCCTTC 505
 Qy 478 ATGTTTAAATGTCTGTAGACAGCTCAACAGATGTGGGAAAACATGCTAACCTGAGT 537
 Db 506 ATGTTTAAATGTCTGTAGACAGCTCAACAGATGTGGGAAAACATGCTAACCTGAGT 565
 Qy 538 GATGACCAAAATATAGTTGCAACAGGCAAGAGTTCTGATTTGGTATGTGATATCA 597
 Db 566 GATGACCAAAATATAGTTGCAACAGGCAAGAGTTCTGATTTGGTATGTGATATCA 625
 Qy 598 AGAGCTGAGCTGCTTTTAAAGAGTTGCAATGCAATTTGGCAGAGTACCAG 657
 Db 626 AGAGCTGAGCTGCTTTTAAAGAGTTGCAATGCAATTTGGCAGAGTACCAG 685
 Qy 658 AATTCATCCGAAGATGCCAAAATGTCTTTCATCTTCAACATCTGAGTCTGGGT 717
 Db 686 AATTCATCCGAAGATGCCAAAATGTCTTTCATCTTCAACATCTGAGTCTGGGT 745
 Qy 718 TTTTACCTGACAGACAGAGGGAGCTTTATCTAAGCTATTTGACAGAGAAAGAGGGGT 777
 Db 746 TTTTACCTGACAGACAGAGGGAGCTTTATCTAAGCTATTTGACAGAGAAAGAGGGGT 805
 Qy 778 ATCAGTTTGTCTGAGTCTTGGTATTTATCTTCTCCATCAGTTTGAAGTATG 837
 Db 806 ATCAGTTTGTCTGAGTCTTGGTATTTATCTTCTCCATCAGTTTGAAGTATG 865
 Qy 838 CTGTGAGTGTGTGTGTATTAACAATGCGGTGCCAGTTTGTCTTGGTATG 897
 Db 866 CTGTGAGTGTGTGTGTATTAACAATGCGGTGCCAGTTTGTCTTGGTATG 925
 Qy 898 ACACCTGGTATATACAGCATTCAGCTTGCAGTCCAGGTGAGAGATAGATTTAGA 957
 Db 926 ACACCTGGTATATACAGCATTCAGCTTGCAGTCCAGGTGAGAGATAGATTTAGA 985
 Qy 958 ATAGAAATGAACAAAGCAGATATGATGACAGTATGCTGCTATAGACTCAGCTGAT 1017
 Db 986 ATAGAAATGAACAAAGCAGATATGATGACAGTATGCTGCTATAGACTCAGCTGAT 1045
 Qy 1018 TATGAAACTGTGAAGTATTTAATATGAAAGATATGAAAGACAGAGATATGATGATTT 1077
 Db 1046 TATGAAACTGTGAAGTATTTAATATGAAAGATATGAAAGACAGAGATATGATGATTT 1105
 Qy 1078 TTGAAAGCTATGAGACGCTTCACTGAAAGTACCTTACTCTGGCTATGCTGAATCTTT 1137
 Db 1106 TTGAAAGCTATGAGACGCTTCACTGAAAGTACCTTACTCTGGCTATGCTGAATCTTT 1165
 Qy 1138 GGTCAAAAGTGTCTATTTTCACTGTGCGTTTAAACAGTATATGCTGCGCCAGTCAAGGA 1197
 Db 1166 GGTCAAAAGTGTCTATTTTCACTGTGCGTTTAAACAGTATATGCTGCGCCAGTCAAGGA 1225
 Qy 1198 ATTGTGCGAGGTACCTTACTGTGAGATCTAGTAATGTGAAATGCACTGCTTTTTCAG 1257
 Db 1226 ATTGTGCGAGGTACCTTACTGTGAGATCTAGTAATGTGAAATGCACTGCTTTTTCAG 1285
 Qy 1258 CTTTCAATTAACCCCTGACCTTCTGGAACCTGATATATGAGACATAGCAAGCACTCATTA 1317
 Db 1286 CTTTCAATTAACCCCTGACCTTCTGGAACCTGATATATGAGACATAGCAAGCACTCATTA 1345
 Qy 1318 GATATGAACACTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1377
 Db 1346 GATATGAACACTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1405
 Qy 1378 GCATCTCCCTTCAAGATCACACAGACAGCTACCGTGGCCTTTGATATGTGCATTTT 1437

Db 1406 GCATCTCCCTTCAAGATCACACAGACAGCTACCGTGGCCTTTGATTAATGTGCATTTT 1465
 Qy 1438 GAATATCATTTAGAGGCCAGAAAGTCTTATGTGATATCTTTGAAGTCCCTGCAGAAAG 1497
 Db 1466 GAATATCATTTAGAGGCCAGAAAGTCTTATGTGATATCTTTGAAGTCCCTGCAGAAAG 1525
 Qy 1498 AAAGTGCCATTTGTAGAGGTAGTGGGTCAAGGAAAAGCACATATGTAGAGCTATTTATTT 1557
 Db 1526 AAAGTGCCATTTGTAGAGGTAGTGGGTCAAGGAAAAGCACATATGTAGAGCTATTTATTT 1585
 Qy 1558 CGCTTCATAGACCTCAAAAAGGTAGATTTATCTTCTGTGTCAAAATATATACAGATGTG 1617
 Db 1586 CGCTTCATAGACCTCAAAAAGGTAGATTTATCTTCTGTGTCAAAATATATACAGATGTG 1645
 Qy 1618 AGCCTGAAAAGCCTTGCAGAGGACAGTGGAGTGGTACTCAGATGTGCTCTTCCAT 1677
 Db 1646 AGCCTGAAAAGCCTTGCAGAGGACAGTGGAGTGGTACTCAGATGTGCTCTTCCAT 1705
 Qy 1678 AATATCTATTTATTAACAACCTTTATATGAAACATCATGCTCTTCACTGAGAAAGTGTAT 1737
 Db 1706 AATATCTATTTATTAACAACCTTTATATGAAACATCATGCTCTTCACTGAGAAAGTGTAT 1765
 Qy 1738 GCAATGCAAAATTAAGTGTGATCTTATGATGCAATTTCTTGAATGCCATATGATATGAC 1797
 Db 1766 GCAATGCAAAATTAAGTGTGATCTTATGATGCAATTTCTTGAATGCCATATGATATGAC 1825
 Qy 1798 ACCCAATGAGGGGAAACGAGGACCTCAAGCTTTCAGAGAGAAAGAAAGAGTATGCAATT 1857
 Db 1826 ACCCAATGAGGGGAAACGAGGACCTCAAGCTTTCAGAGAGAAAGAAAGAGTATGCAATT 1885
 Qy 1858 GCAGAGCCATTTTGAAGAACCCCGCAGTATATCTATATGATGAGTACTTCACTGCTTA 1917
 Db 1886 GCAGAGCCATTTTGAAGAACCCCGCAGTATATCTATATGATGAGTACTTCACTGCTTA 1945
 Qy 1918 GATTCATATCTGAAGAGACTATTTCTGTGTCATGAAAGATGTGTCAAAACAGAACT 1977
 Db 1946 GATTCATATCTGAAGAGACTATTTCTGTGTCATGAAAGATGTGTCAAAACAGAACT 2005
 Qy 1978 TCTATTTTCAATGACACAGATTTGTCAACAGTGGTATGATCAGATGAATCATATGCTTGT 2037
 Db 2006 TCTATTTTCAATGACACAGATTTGTCAACAGTGGTATGATCAGATGAATCATATGCTTGT 2065
 Qy 2038 GATCAGGGTAAAGTAAAGCGGAGTGTACCCACCATGTTTCTGCTAACCTCATAGT 2097
 Db 2066 GATCAGGGTAAAGTAAAGCGGAGTGTACCCACCATGTTTCTGCTAACCTCATAGT 2125
 Qy 2098 ATCTATTTCAAGAAATGTGGCATACACAGAGACCGGTGTGCAGAAACCATGATACCCCAA 2157
 Db 2126 ATCTATTTCAAGAAATGTGGCATACACAGAGACCGGTGTGCAGAAACCATGATACCCCAA 2185
 Qy 2158 TGGGAGGC 2165
 Db 2186 TGGGAGGC 2193

RESULT 8
 MMU43892 2684 bp mRNA linear ROD 28-Apr-1997
 LOCUS
 DEFINITION Mus musculus ABC transporter-7 mRNA, partial cds.
 ACCESSION U43892
 VERSION U43892.1 GI:1167981
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2684)
 AUTHORS Savary,S., Allikmets,R., Denizot,F., Luciani,M.F., Mattei,M.G., Dean,M. and Chimi,G.
 TITLE Isolation and chromosomal mapping of a novel ATP-binding cassette transporter conserved in mouse and human
 JOURNAL Genomics 41 (2), 275-278 (1997)
 MEDLINE 9728528

PUBMED 9143506
REFERENCE 2 (bases 1 to 2684)
AUTHORS Savary,S., Denizot,F., Matrei,M.G. and Chimini,G.
TITLE Direct Submission
JOURNAL Submitted (26-DEC-1995) Stephane Savary, Centre d'Immunologie de
Marseille Luminy, Parc Scientifique et Technologique de Luminy,
Case 906, Marseille Cedex 9, 13288, France

FEATURES
source
1..2684
/organism="Mus musculus"
/strain="DBA/2"
/db_xref="taxon:10090"
/chromosome="X"
/map="XC-D"
/clone="lambda ABC7-3b"
/cell_line="P388D1"
/cell_type="monocyte/macrophage"
<1..2085
/codon_start=1
/product="ABC transporter-7"
/protein_id="AAC5152.1"
/db_xref="gi:1167982"
/translation="ESLRNTTQQRKGNKNSROLLDATKALQTWPLEIKRTGWHGAG
GLHTDPKGLDQVDRKTIKAMLSYVPEDEPDLRARAISLGFGLGKAKNIVPEM
FKYAVDSLQMSGMMLNSDAPNTVAATAVLIGVSRGAAPFNEVRNAVGVKA
ONSIRIKNTNFIHLNLDGFLHSRGTGALSKAIDRTGRISEVLSALVNLPIVF
EMMLVSVLYKCGAOFALVTIGTATFAVTAFTVRRTFRLEMKKANDAGNAAI
DSLNTETVKYFNNEKYPARQYDGEIKETYSIKSTSTAMLNPGMTAFVGLTAT
MYLASQIVAAALTVGDLVWNGLLFQSLPLNPLGTVYRTRQALLDMNTLFTLKY
DTRIKDKWAPPLQITPQTAVAFDNVHEFIEQGLRANVGVSEVPAPKVAITFLYK
LQKSTVRLFRFEPOKSIYLAGQNLQDVLSLRVAVVPODVALFNNTIYNNL
LYGNNINSPREVVAVAKLAGLIDALRMPHYDQVQVGRKDEGEORVAIRAIL
KNPPIYLDKATSSLDSTIEETILGAMNDVHKRDIPIARLSIVVDADIEIYLSOG
KVAREGTHYGLANSSSIYSEMHMTQSNRVQNSLGDAKKESLSKEBEKKQI0E01
VNSYKGCNCSC"

BASE COUNT 818 a 502 c 581 g 782 t 1 others
ORIGIN

Query Match 78.2%; Score 1833; DB 10; Length 2684;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1941; Conservative 1; Mismatches 181; Indels 0; Gaps 0;

Qy 175 GAGTCATTAAAAAGTATCATGAGGAGATTGGGAAAAAGCAATTGACGAGCTTTCTTA 234
Db 1 GAATCATTTAAAGAAATACACACAGCAGAGATGGGAAAAAGACAATCCGACAGTTACTTA 60
Qy 225 GATGCTGCAAAAGCTCTCCAGGTATGGCCACTGATAGAAAAAGACATGTTGGCATGCT 294
Db 61 GATGCTACAAAGCTCTCCAGCATGGCCATTGATAGAAAAAGACATGTTGGCATGGG 120
Qy 295 CATGCAGAGAGAGACTCCACACAGACCMAAAGAGGTAAAGATGTGATCTGGG 354
Db 121 CACGCAAGAGAGACTCCACACAGACCMAAAGAGGTAAAGATGTGATCTGTA 180
Qy 355 AAAATCATAAAAAGCATGCTTTCTTATGTGTGCCCCMAAGACAGGCGAGATCTACAGCT 414
Db 181 AAAATCATAAAAAGCATGCTTTCTTATGTGTGCCCCMAAGATAGGCGAGATCTACAGGCG 240
Qy 415 AGAGTTGCCATTTCCGTGGGATTTTGGGTGTGCAAAAGCCATGAAATTTGTGTTCCC 474
Db 241 AGAGTTGCCATCTCCCTGGGATTTTGGGTGTGCAAAAGCCATGAAATTTGTGTTCCCT 300
Qy 475 TTTCATGTTTAAATATGCTGTAGACGCTCAACCAAGATGGGGAACATGCTGAACCTG 534
Db 301 TTTCATGTTTAAATATGCTGTAGACGCTCAACCAAGATGGGGAACATGCTGAACCTG 360
Qy 535 AGTGAATGCAAAATACAGTTGCAACCATGAGCAACAGCAGTTCTGATTTGGCTATGGTGA 594
Db 361 AGTGAATGCAAAATACAGTTGCAACCATGAGCAACAGCAGTTCTGATTTGGCTATGGTGA 420
Qy 595 TCAAGAGCTGAGAGCTGCTTTTAAACGAAGTTCCAAATGCGATTTTGGCAAGTAGGCC 654
Db 421 TCAAGAGCTGAGAGCTGCTTTTAAACGAAGTTCCAAATGCGATTTTGGCAAGTAGGCA 480

Qy 655 CAGATTCATCCGAGAGATAGCCAAAAATGTCTTCTCCATCTTCAACAACCTGGATCTG 714
Db 481 CAATATTCATACGAGAGATAGCCAAAAATGTATTTCTCCATCTTCAACAACCTGGATCTG 540
Qy 715 GGTTTTCACTGACGACAGACAGCGGAGCTTTATCTAAGGCTATTGACAGAGAAACAAG 774
Db 541 GGTTTTCACTGACGACAGACAGCGGAGCTTTATCTAAGGCTATTGACAGAGAAACAAG 600
Qy 775 GGATACGTTTGTGCTGAGAGCTTTGGTATTTATCTCTCTCCATCATATGTTGAAGTG 834
Db 601 GGATACGTTTGTGCTGAGAGCTTTGGTATTTATCTCTCTCCATCATATGTTGAAGTG 660
Qy 835 ATGCTTGTCACTGAGTGTGTTTGTATTAACAATCCGTCGCCAGTTTGGTTGTAACCTT 894
Db 661 ATGCTTGTCACTGAGTGTGTTTGTATTAACAATGCTGTCGCCAGTTTGGTATGTAACCTTA 720
Qy 895 GGAACACTTGGTACATACACAGCATTTCAAGTTCAAGTCCACAGCGTGGAGAACTAGATT 954
Db 721 GGAACACTTGGTACATATACAGCATTTCAAGTTCAAGTCCACAGCGTGGAGAACTAGATT 780
Qy 955 AGAATAGAAATGAACAAAGCAGATATAGTACAGTAAATGCTGTATAGACTCACTGCTG 1034
Db 781 AGAATAGAAATGAACAAAGCAGATATAGTACAGTAAATGCTGTATAGACTCACTGCTG 840
Qy 1015 AATTATGAACCTGGAAGTATTTTAAATGAAGATATGAAGACAGAGATATAGTGA 1074
Db 841 AATTATGAACCTGGAAGTATTTTAAATGAAGATATGAAGACAAAGATATAGTGA 900
Qy 1075 TTTTGAAGAGATATGAGAGCTGCTTCAATGAAGATCTTCACTCTGGCTATGCTGAA 1134
Db 901 TTTTGAAGAGATATGAGAGCTGCTTCAATGAAGATCTTCACTCTGGCTATGCTGAA 960
Qy 1135 TTTTGAAGAGTATTTTGAAGTGTGTTTGAAGTGTGTTTGAAGTGTGTTGAGTGTG 1194
Db 961 TTTTGAAGAGTATTTTGAAGTGTGTTTGAAGTGTGTTTGAAGTGTGTTGAGTGTG 1020
Qy 1195 GGAATGTGAGAGTATCCCTTACTGTTGAGATCTAGTAATGTGAATGCACTGCTTTT 1254
Db 1021 GGAATGTGAGAGTATCCCTTACTGTTGAGATCTAGTAATGTGAATGCACTGCTTTT 1080
Qy 1255 CAGCTTTCATTAACCTCTGAACTTTTCTGGGAACTGATATAGAGACTAGACAACTC 1314
Db 1081 CAGCTTTCATTAACCTCTGAACTTTTCTGGGAACTGATATAGAGACTAGACAACTC 1140
Qy 1315 ATAGATATGAACCTGTTTACTCTCTCAAGGTAGCAACCAATTAAGACAAAGTG 1374
Db 1141 ATAGATATGAACCTGTTTACTCTCTCAAGGTAGCAACCAATTAAGACAAAGTG 1200
Qy 1375 ATGCACTCTCCCTTCAAGATCACACAGACAGACTACCGTGGCTTTGATATGTGAT 1434
Db 1201 ATGCACTCTCCCTTCAAAATTAACACACAGACAGACTACCGTGGCTTTGATATGTGAT 1280
Qy 1435 TTTGAATACATTAAGAGGCGAGAAAGTCTTATGTGAATATCTTTGAAGTCTGCAAGA 1494
Db 1261 TTTGAATACATTAAGAGGCGAGAAAGTCTTATGTGAATATCTTTGAAGTCTGCAAGA 1320
Qy 1495 AAGAAAGTGGCAATTTGATAGAGTATGAGGTGCAAGGAAACCAATATAGAGCTATTA 1554
Db 1321 AAGAAAGTGGCAATTTGATAGAGTATGAGGTGCAAGGAAACCAATATAGAGCTATG 1380
Qy 1555 TTTGCTTCTATGAGAGCTCAAAAGGTTAGCAATTTATCTTGTGTCAAAATATACAGAT 1614
Db 1381 TTTGCTTCTATGAGAGCTCAAAAGGTTAGCAATTTATCTTGTGTCAAAATATACAGAT 1440
Qy 1615 GTGAGCTTGAAGAGCTTTGAGAGGCGAGTGGAGTGTACTCAAGATGCTGCTCTTC 1674
Db 1441 GTGAGCTTGAAGAGCTTTGAGAGGCGAGTGGAGTGTACTCAAGATGCTGCTCTTC 1500
Qy 1675 CATATATATATTTATTAACAACCTCTTATATAGAAACATACGCTTCACTGAGAGAGTG 1734
Db 1501 CATATATATATTTATTAACAACCTCTTATATAGAAACATACGCTTCACTGAGAGAGTA 1560

Qy	774	GGGATTCAGTTTTGCTCTGAGTCTTTGGATTTAACTCTTCCACATCAGTTTGAAGT	833
Db	1118	GGGCAATTAACTTTGTGTCTCTCCGCATGTGCTTCAACATTTGGCCCATCTATTGGAGCT	1177
Qy	834	GATCGTTGTCAGTGTGTGTTTTGTATTACAATATGCGGTGCCAGTTTGGTTGTGAACCTT	893
Db	1178	GGCCCTGCTGTCCAGTATCCTGGGAATGAAAGTGTGCGCTGGCTTCCGTGTGTCAAGAT	1237
Qy	894	TGGAACCTTGATACATACACAGCATTTCAAGTTGCATGCACAGGTGAGAACTAGATT	953
Db	1238	GGGCTGCGTTGGCAATATACCTGCTCAACTCGATGTGTGACCCAGTGGCCACCCAGTT	1297
Qy	954	TAGAATAGAAAAGAACAAACAGATTAATGATGAGTTATGTGTATATAGCTCAGCTCT	1013
Db	1298	CCGGTATTTATAGAACAGCGCGAGAACAGAGCGCGCAACAAAGCCGTAGACTGTGAT	1357
Qy	1014	GAATTAATGAACCTGTGAAGTATTTTAAATATGAAGAAATATGAAGACAGATATGATGG	1073
Db	1358	CAACTACGAGCGGTGAAGTACTTCAACACGAAAGTACAGAGGGGTGTCTACACAGA	1417
Qy	1074	ATTTTGAAGCGTATGAGACTGTCTTATTGAAAAGTACCTTACTCTGGCTATGTGA	1133
Db	1418	GTGTGTGAAGAAAGTATGAGCGCGCAGCTTGAAGACTAGCTCAGTCTGTCTGTCAA	1477
Qy	1134	CTTTGTGTCAAGTGTCTATTTTTCAGTGTCCGTTTAAAGCTATTAATGTGTCTCCAGTCA	1193
Db	1478	CTTTGGCCGAGAACCCATTTTCAGACAGCGCATGATCTGATTTATGTATTTGGCCGCCAA	1537
Qy	1194	GGGAATTTGTGCAGTACCCCTTACTGTGTGAGATCTAGTAATGTGAATGACACTGCTTTT	1253
Db	1538	GGAGATTTGCCAAGGCCAATGACGTGTGGAATTTGTGTGATGTCTACAGCCCTGTCTT	1597
Qy	1254	TCAGCTTTCATTACCCCTGGAACCTTTCGGAACTGTATATAGAGACTAGACAAGCACT	1313
Db	1598	CCAGCTTCGATCCCCCTCGGTTTTCTGGCAGTGTATTATCGTAGGTGCACAGGCTTT	1657
Qy	1314	CATGATATATGAACACTTGTTTTACTTACTCAAGGTAGACACCCCAATTTAAAGCAAAAGT	1373
Db	1658	GCTGACATGTGGGGCCATGTGTTCAGTTGAAGACGTGACAGATGACATTTCAACAGCGCC	1717
Qy	1374	GATGCAATCTCCCTTCAGATCAACCCAGACAGCACTACCGTGGCCTTGTAAATATGTCA	1433
Db	1718	CAATCCCAAGCCCTGTTCTGTGACACACTAATCTATCAATTGAGTTCCGACAGTATAG	1777
Qy	1434	TTTTGAATATCATTTGAGGCCAGAAAGTCTTATGTGAATATCTTTGAAGTCCCTGCAGG	1493
Db	1778	CTTGAAGTACGAACTGTGGCAAGCCCAATTTTCCGGAGCTCATGTTCAACATATCCCGCGG	1837
Qy	1494	AAAGAAATGTGCCATTTTATGAGAGTATGTGGGTCAAGGAAAAGACAAATATGTAGGCTATT	1553
Db	1838	TAAAAACCTGCCATGTGTGGCGGCTGTGCTCAGGAAAAATCGTGCATGTGTGCGCTTCT	1897
Qy	1554	ATTTGCTTTCATATAGGCTCAAAAAGGTATACATTTATCTTGCTGTCAAAATATACAGA	1613
Db	1898	CTTCCGCTTTCTTTAGCCAAACTGTGTTAAAGTTTGATTCGTGTGCGCAGAGATATAGCGC	1957
Qy	1614	TGTGAGCTTGAAAAGCCTTTCGAGAGGCAGTGGAGTGTAACTCAGAGTGTCTCTT	1673
Db	1958	CGTGAATTTTGAAGAGCCTGCGCAAGTTATTTGCAAGTGTGTGCGCAGAGATTTGGTGTATT	2017
Qy	1674	CCATATATCTATTATTTACAACTCTTATATGAAAAATCATGAGTCTTCACTTGAGGAAGT	1733
Db	2018	CCACAACACATATGAGCAATATATCCACTATGTGCAACTGTCCAAAGTCCCAAGCGGAAGT	2077
Qy	1734	GTATGCAATGGCAAAATATAGTGTGACTTCAATGATGCAATTTCTTCAATGCCCATGTAGTA	1793
Db	2078	GGAGATCTCTCCCGCATAGCTGATTTTGCAATGACTCCATATATAGCTGTGCGCAGGACAGTA	2137
Qy	1794	TGACACCCCAATAGGGGAAACAGAGACTCAAGCTTTTCAGAGAGGAAAAAGCAAGTATAGC	1853
Db	2138	CTCCACTAGGTATGAGCCAACTGTGATTTAAAGTTATCCGGTGGGAGAAACAGACGGGTTCG	2197
Qy	1854	AAATTCAAGAGCCATTTTGAAGACCCCCCAAGTCATATCTTATGATGAAGTACTTCAATC	1913

QY	307	GGAGTCCACAGACCCAAAAGAGCGTTAAAGATGTTGTAATCTCGGAAATCATATAAA	366
Db	2198	CATGCGTAGGGCCACTACTGGAAGAACACCCCATCTTGTATCTTTACACAGCGTTAACAGAG	2257
QY	1914	GTTAGATTGCATTACTGAGAGAGACTATTCTTGGTCCATGAGAGATGTGTCMAACAG	1973
Db	2258	TCTGGACTCGATCACCCGAGCGTAACATTCTGACAGCTCTGACCCCGCTACTCCAGACG	2317
QY	1974	AACCTTATTTTTCATTGACACAGATTGTCAACAGTGGTTGATGCAGATGAATCATGT	2033
Db	2318	CATTAGCATTTTGACATGCGCCACCGCTCTCGACAGTCAAGATCGGACGAGATTCTGCT	2377
QY	2034	CTTGATCAGAGGTAAAGGTAGCCGAAACGTCGTACCACCATGTTTGGCT	2081
Db	2378	GCTCGAAGACGACGCTGTAGGTGAACGTGGACCCCATTCGAGAGCTGCT	2425
RESULT	10		
LOCUS	AF287699	2187 bp	mRNA linear
DEFINITION	Arabidopsis thaliana half-molecule ABC transporter ATM3 mRNA,		PLN 02-SEP-2000
ACCESSION	AF287699		
VERSION	AF287699.1		
KEYWORDS	GI:9964120		
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 2187)		
TITLE	Sanchez-Fernandez, R., Mari, S., Dancis, A. and Rea, P.A.		
JOURNAL	Functional half-molecule ABC transporters from Arabidopsis: the ATM family		
AUTHORS	2 (bases 1 to 2187)		
TITLE	Sanchez-Fernandez, R., Mari, S., Dancis, A. and Rea, P.A.		
JOURNAL	Direct Submission		
TITLE	Submitted (17-JUL-2000) Plant Science Institute, Department of Biology, University of Pennsylvania, 3800 Hamilton Walk, Philadelphia, PA 19104-6018, USA		
FEATURES	Location/Qualifiers		
source	1..2187		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/note="ecotype: Col-0"		
	1..2187		
	/codon_start=1		
	/product="half-molecule ABC transporter ATM3"		
	/protein_id="AA09829.1"		
	/db_xref="GI:9964121"		
	/translation="WSRGRKVRAPGILLRLRVNLQPPKIPSPSYLSRSDYRLHNGS		
	/transl_except="KTKRIKTSSDSDAMADMKILRTLAICYLMMRDPERFRVIALGELVAGKLVNVOV		
	PTFKLVADMLASATGASLTTPNATPLVFAFPAAVLISYGALARGSSAFNEEL		
	RVNVPTTLEISWISGLAIKFKGAAPMTISLSGSIIVPLATVQRTKRMKMK		
	DNDSTRAIDSLINIVETVKYFNNEGYAEKXDOCLKYEADLQTSLSLAFNFGQST		
	IFNSALSTAWLCSQIINGQMTVGDVVMNGLFQSLPLNGLSVYRETIQSLVVM		
	KSMQLLEKSDITNTSDAKPLVKGNIIFENYHFSYLERKILDISFVVPAGKSSV		
	ALVETSGSKSDITLRMLFRPDDTSGNIRIDGDIKEVRLDSSLSIGVODTVLFR		
	DTIHNIHNGRLSATREVEYEAARRAIHETISNPDKSVTIYVERGLKLSGGRKOR		
	IVLTFEISPRILICEDATSLDSTAEILNALKALASRTSIFIAHRLTTMAQCEDE		
	ALAVENGVQEGHDELLGSGSYADLWLTQONSSVMDLAATKLE"		
BASE COUNT	598 a 479 c 483 g 627 t		
ORIGIN			
Query Match	23.2%	Score 543.6;	DB 8; Length 2187;
Best Local Similarity	57.5%;	Pred. No. 1.5e-122;	
Matches 1069;	Conservative 0;	Mismatches 764;	Indels 27; Gaps 4;
QY	307	GGAGTCCACAGACCCAAAAGAGCGTTAAAGATGTTGTAATCTCGGAAATCATATAAA	366
Db	316	GAAATCAAGACCACTTCGCGGATTCAGACTCAGCTTATGCTGACATGAAGATTCTCCGT	375

Qy 367 GCAATGCTTTCTTATGTGTGCCCCAAAGACAGGCAATCTAGAGCTAGAGCTTCCATT 426
 Db 376 ACGCTTGTGCTATTTTGTGATGAGACAACTCCGAATTCGGTTACGGGCTCATCCCT 435
 Qy 427 TCGGGGATTTTGGGTGGTCAAGGCAATGATATTTGGTCCCTTCACTGTTTAA 486
 Db 436 GCTTTGGGTTTCTGTGGGTGCTAAGGCTTGAAATGTCAAATCCCTTCTCTTCAA 495
 Qy 487 TATGCTGTAGACAGCT-----CAACGATGTCCGGAAACATGCTGAACCTGAGTG 538
 Db 496 TTAGCTGTGATGTGCTTGTCTACCTACCGCTACTGAGGCTCACTTACCACTTCGCT 555
 Qy 539 ATG-----CACCAATACAGTTGCAACCATGGAACAGAGCTTCTGATTTGGCTAT 588
 Db 556 GCTACAAACCAACTCTTCTTACTGTTTGTGCTACTCCGTGCTGTTCTCATTTGATAC 615
 Qy 589 GGATGATAGAAGCTGAGCTGCTTTTTTTAAAGAAAGTTTCAAAATGCGATATTTGGCAAG 648
 Db 616 GGGATTTGCTCGGACGGGTTCTTGTCTTTTATGAACTTGTATGCTGTGTCTCTAAG 675
 Qy 649 GTAGCCCAAGATTCATCCGAAGATAGCCAAATGTCTTCTCCATCTTCAACAACCTG 708
 Db 676 GTAGCTCTCCGAACAACTACGATCTGTCTTGAAAGGTGTTCTCATTTTACATGATCTT 735
 Qy 709 GATCTGGGTTTTCACCTGAGACAGACGAGGAGCTTATCTAAAGCTATTGACAGAGA 768
 Db 736 GACCTTCGTTACACCTAAGTCGAGAAACAGCGGCTTGAAACAGGATTTATGATGTGT 795
 Qy 769 ACAAGGGATACAGTTTGTCTCTGAGTCTTGTGATTTAATCTTCTTCCATCATGTTT 828
 Db 796 AGCGGTGAATTAATTTTATCTCTACACTATGTGTTTCAATGTTGTCTTCAATATAG 855
 Qy 829 GAAGTATGCTTGTCAAGTGTGTTTGTATTAACAATGCGGACCAAGTTTCTTGGTA 888
 Db 856 GAGATATCTATGATATGATAGTATCTGCAATACAGTTTGAAGCTGCTTGTCCGATTT 915
 Qy 889 ACCCTTGAAACATTTGTGATACACAGCATTTCAAGTTGAGTCAACGCTGAGAACT 948
 Db 916 ACTTCTCTCTCGTTGGATCATATATGCTTTTCAATAGCTGTGACACATGTGCAACC 975
 Qy 949 AGATTTGAATAGAAATGAACAAAGCAGATATATGATGAGTAATGCTGCTATAGACTCA 1008
 Db 976 AAGTTTGAAGGCTATGAAACAAAGCTGATATGATGCTACACAAAGGCCATGATTTCT 1035
 Qy 1009 CTGCTGAATTTAGAACTGTGAAGATTTTATATATGAAGAAATATGAAGCAGAGATAT 1068
 Db 1036 CTGATTAATTTATGAGACTCTCAAAATATTTTAAACAATGAGTTATGAGGCTGAAAGAT 1095
 Qy 1069 GATGATTTTGAAGACGTATGAGCTGCTTCAATTGAAATGATCCTTACTCTGGCTATG 1128
 Db 1096 GACCAAGTTCCTGAATAAATATGAAAGATGCTCCCTGCAAACTCAAAAGATCTTGCTTC 1155
 Qy 1129 CTGAACCTTTGGTCAAGTGTCTATTTTCAAGTGTGCGTTTAAACAGCTATATAGTCTGCGC 1188
 Db 1156 TTAAACTTTTGAACAAGATATCATATTTTATGACCTTTATCAACAGCAATGTGTGTGC 1215
 Qy 1189 AGTCAAGGAATGTGGACAGTACCTTACTGTGAGATCTAGTAATGTGCAATGAGCTG 1248
 Db 1216 TCTAGGGTATCATGAATGGCAGATGACATCGGTATTTGGTTATGTGGAATGAGCTT 1275
 Qy 1249 CTTTTCAGCTTTTCAATACCCCTGAACCTTCTGGAACCTGTATATGAGAGACTAGACAA 1308
 Db 1276 CTTTTCAGGCTCTCTTCTCTAAATTTTGGGTAGTGTGTTACGTGAAACAAATTCAG 1335
 Qy 1309 GCACTCTAATATGAACACCTGTTTACTCTACTCAAGCTAGACACCCAAATTAAGAC 1368
 Db 1336 AGCCTTGATGACATGAATCATATGTTCCAGTTGTAGAGGAATCTGACATCACAAT 1395
 Qy 1369 AAAGTATGAGCATCTCCCTTCAAGATCAACACAGACAGTACCGTGGCTTTGATTAAT 1428
 Db 1396 ACAAGTACCGCAAGCTCTT-----GTACTGAAGGTGGAACATTTGTTGAAAT 1449

Qy 1429 GTGCATTTTGAATACATTGAGGGCCAGAAAGTCCCTAGTGAATATCTTTGAAGTCCCT 1488
 Db 1450 GTACACTTCAAGTATCTTCCAGAGAGAAATATTAATGGAATTTCTTGTGTGTAACG 1509
 Qy 1489 GCAGAAAGAAAGTGGCATTGTAGAGTGTGTGGTTCAGGAAAGACAAATAGTAGAG 1548
 Db 1510 GCAGAAAAAAGTGTGGCAATGTGTGTATAGCGGGAGCGGCAAGTCAAAATCTTAGG 1569
 Qy 1549 CTATTAATTTGCTTCTATGAGCTCAAAAAGGGTGAATTTATCTTGTGTCGCAAAATATA 1608
 Db 1570 ATGCTATTTCAAGATTTCTTGCACACAGATTTCCGAAATATTAAGATTTGACGGCAAGATATA 1629
 Qy 1609 CAAGATGTAGCTTGAAGGCTTCCGAGGGCAGTGGAGGTGTACTCAGATGCTGTCTC 1668
 Db 1630 AAGGAAGTGAAGTATGATAGCTCCGAAAGCTCATAGGAAGTGTGTCCGCAAGATATCTGT 1689
 Qy 1669 CTCTTCAATATCTATTTATTTATACACCTCTTATATGAAACATCAGTCTTCACTGAG 1728
 Db 1690 CTCTTCAATGACACAAATTTTCAACATATCATTTATGAGCGTCTCTCAGCTACAGAGAA 1749
 Qy 1729 GAAGTATGAGTGGGCAAAATTAAGCTGAGCTTCAATGATGCAATTTCCGATGCGCAT 1788
 Db 1750 GAGGTGATAGGCTGTCTCGCGCTGTGCAATTCATGAAACATGATTAATTTCCGTGAT 1809
 Qy 1789 GGAATATGACACCCAAAGTAGGGGAACGAGACTCAAGCTTTTCAAGAGAGAAACCAAGA 1848
 Db 1810 AAATATTTCCAGATCGTTGGGGAAGGGGCTCAAGTTGAGTGTGAGAGAAACAAAGA 1869
 Qy 1849 GTAGCAATTTGCAAGAGCCATTTTGAAGAGACCCCACTCATCTATGATGAAGTACT 1908
 Db 1870 GTGGCACTTCTCGGACATTTTGAATCCCTGCGATTCGTGTGTGATGAGGCAACG 1929
 Qy 1909 TCATCGTATTCGATTTACATGAAAGACTATCTTGTGTCATGAAAGATGTGTCATA 1968
 Db 1930 AGTGGCTGACACGACACAGAGGACAGATATGAAACGACCTCAAGGACTAGCGAGT 1989
 Qy 1969 CACGAACCTTCTATTTTTCATTGACACAGATTTGCAACAGTGTGTATGAGATGAATC 2028
 Db 1990 AACGAGCATCTATCTTCAATTTGCCACAGGCTGACCTCGATGCAATGTAGAGATC 2049
 Qy 2029 ATTGTCTTGATCAGGGTAAAGTAAAGTAAAGTGTATCCCAATGAGTTGTGCTTGAAC 2088
 Db 2050 GTAGTACTGGAACCGGGAAGGTGTGGAACAAAGGCCCCACACAGCAGCTATTGGGAAAA 2109
 Qy 2089 CTTCATATGATCTTATTCAGAAATGTGSCATACACAGACGCGGTGCGAAGCAATGAT 2148
 Db 2110 TC---AGGGAGATATGCGCAGCTATGACACAAACAGACGCGTGAATATGCTTGAAT 2166

RESULT 11
 AF360334
 LOCUS
 DEFINITION
 Arabidopsis thaliana putative ABC transporter protein (At5g58270)
 mRNA, complete cds.
 ACCESSION
 AF360334.1 GI:13430843
 VERSION
 FLI CDNA.
 KEYWORDS
 SOURCE
 Arabidopsis thaliana.
 ORGANISM
 Arabidopsis thaliana.
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 2373)
 REFERENCE
 AUTHORS
 Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banb, J., Chung, M.K.,
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
 Cariminci, P., Chen, H., Cheuk, R., Hayaishi, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
 Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,
 Davis, R.M., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 TITLE
 Journal Unpublished
 REFERENCE
 2 (bases 1 to 2373)


```

Db 1473 GTACCTTACGTTACCTTCCAGAGAGAAAATATTAGATGAAATTTCTTTGTGTAACG 1532
Qy 1489 GCAGGAGAGAAAGTGGCCATTGTAGAGGTAGTGGGTGAGGAGAAAGCAATAGTGAG 1548
Db 1533 GCAGGAGAGAAAGTGGCAATGTTGTACTAGCGGAGAGCGGAGTCAAAATCTTAG 1592
Qy 1549 CTATTATTCGCTTCTATAGACCTCAAAAAGGTAGCATTTATCTTGTGTCGTAATATA 1608
Db 1593 ATGCTATTCACATTTCTTGCACACAGATTCGGAATATATAGAGATGACGGCAAGATATA 1652
Qy 1609 CAAGATGAGCCTGAGAAAGCCTTCGAGGAGAGGAGGTGATCCTCAGAGATGCTGC 1668
Db 1653 AAGGAGTGAACCTAGATAGCCTCCGAAGCTCCATAGAGATGTGCGCAAGATATGTT 1712
Qy 1669 CTCTCCATTAATTAATTAATTAACACCTTTATATGAAACATCAGTGTCTTCACTGAG 1728
Db 1713 CTCTTCATGACACAAATTTTCAACACATATATGACGCTCTCTCAGCTACAGAGAA 1772
Qy 1729 GAAGTATGACAGTGGCAAAATTAGCTGACTTCATGATGCAATTTCTTGAATGCAAT 1788
Db 1773 GAGGTATGAGGCTGCTCGGCGTGTCTCAATTCATGAACATCAGTAATTTCCCTGAT 1832
Qy 1789 GGATATGACACCCAAAGTAGAGGAAACGAGACTCAAGCTTTCAAGAGAGAAAGCAAGA 1848
Db 1833 AAATATTCACAGATGTTGGGAGAAAGGGGCTCAAGTTGAGTGGAGAGAAACAAAGA 1892
Qy 1849 GTACCAATTCGAAAGGACATTTTGAAGAGACCCCACTACTCTATGATGAAGTACT 1908
Db 1893 GTGGCACTGTGCTGGACATTTTGAATCCCTCGATCTCTGTGTGTGATGAGCAAG 1952
Qy 1909 TCATGTTAGTTCGATTACTGAAGAAGATATCTTGTGTGTCATGAGATGTGTCAA 1968
Db 1953 AGTGCCTGGACAGACAAACAGAGGAGAGATGACTCAAGGAGCACTAGCGAGT 2012
Qy 1969 CACGAACTTCTATTTTATTTGACACAGATTTGCAACAGTGTGATGACAGATGAATC 2028
Db 2013 AACGAGATATCTATCTTATTTGCTCCAGCGCTGACACTGTGATGATGAGAGATC 2072
Qy 2029 ATTTCTTGTGATTCAGAGGTAGAGTGCAGAGTGTGACCCACATGTTTCTTGTCTAAC 2088
Db 2073 GTAGTACTGGAAGAGAGAGAGTGTGAGAACAGGCCCCACGACAGACTTTTGGAGAAA 2132
Qy 2089 CCTCATATGATATTCAGAAATGTGSCATACAGAGAGCCGTGTGACAGAACATGAT 2148
Db 2133 TC--AGGAGATATGCGCAGCTATGACACACAAAACAGAGGTGTGATGCTTGAT 2189

RESULT 12
AF287697 2037 bp mRNA linear PLN-02-SEP-2000
LOCUS Arabidopsis thaliana half-molecule ABC transporter ATML mRNA,
DEFINITION complete cde.
ACCESSION AF287697
VERSION AF287697.1 GI:9964116
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 2037)
AUTHORS Sanchez-Fernandez, R., Mari, S., Dancis, A. and Rea, P. A.
TITLE Functional half-molecule ABC transporters from Arabidopsis: the ATM
family
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2037)
JOURNAL Sanchez-Fernandez, R., Mari, S., Dancis, A. and Rea, P. A.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2000) Plant Science Institute, Department of
Biology, University of Pennsylvania, 3800 Hamilton Walk,
Philadelphia, PA 19104-6018, USA
FEATURES
Location/Qualifiers

```

```

source 1. 2037
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/notes="ecotype: Col-0"
CDS
1..2037
/codon_start=1
/product="half-molecule ABC transporter ATML"
/protein_id="AG09827.1"
/db_xref="GI:9964117"
/translation="MMRGRFLPLRSALYRLRSDHHHSFRIKRSIORPAINA
FTDPSPSPVNARVFPSTSTSTNOOTKASSKILRTISVLMKNDLRF
VIAALCLIGAKTNQVPLPFLSLIDLSYSSSTTBSNPYLLAFAFPSSVLRF
GIARSGSSAPNELRTAVFSKSLRTIRSVSRKVLSHLDLRYHLAREGALNRLID
RGSRAINTLSAMFNVPTLIEISWVTGILANFQPVRLITSLVSGYIAFTLVVT
QYRTFKRKAMQNDNDASTRAIDSLNVEYVKKFNNEYEARXDDLFGSEDAALQT
OKSLAFLDGOSFIFSTALSTSVLCSOGIMENMTVDLVMNGLLFOLSLPLFLG
GYRETVQGLVDMKSLFOLLERSDIDGMDTEKLPYLVRGSLSPENVHPSVLPGR
KIIDGISEFVPRKSVIATVSSSGSTLIRMFRRFPDSDGNVRIDGODIKRYTLES
LNSCIGVPODVLVENDTLFHNHYGNLSATREEVYDAARAVIHDTIMKFPKYSTA
VBSRGMLSGEKORVALARAFLKSPAILLDEATNALDSKTEAEIMKTFRLASNRT
CIFIAHRLTTAMQDSEIIVMEKRVKEKTHQVILBSGRYAKLTQMNSTLEV"
BASE COUNT 576 a 407 c 439 g 615 t
ORIGIN
Query Match 23.2%; Score 543.4; DB 8; Length 2037;
Best Local Similarity 57.2%; Pred. No. 1.7e-122;
Matches 1026; Conservative 0; Mismatches 761; Indels 6; Gaps 2;
Qy 350 CTCGGAATATCATAAAGCAATGCTTCTATGTGCGCCAAAGACAGCGCATATCTAC 409
Db 239 CTAAAGAGATCTCCGACAGATTTCAAGCTATCATGATGAAGAACACCAAGACTCC 298
Qy 410 GAGCTAGATGTCATTTCCCTGGATTTTGGGTGTGCAAGAGCCATGAATATGTG 469
Db 299 GATTGAGATCATGCGCTTTAGCGTCTCATAGAGCTAAGTTTGAATTTCAAG 358
Qy 470 TTCCCTTCAATGTTAAATATGCTGTAGACAGCTCAAC--AGATTCGGAGAAACATGC 526
Db 359 TTCCCTTCTCTTCAATTTATTCATGATTTACTCTCTTCATATTCATCTTCAACAAATCA 418
Qy 527 TGAACTGATGATGACCAATAACATGTCACACATGCGAACAGCGATTTCTGATGGCT 586
Db 419 CTGATTCAAATCCCTATCTGCTGCTGCTTGCCTATCTCTTCTCGTTCTCAATGGCT 478
Qy 587 ATGGTATCATCAAGAGCTGAGCTGCTTTTTTTAAAGAAAGTTCGAAATAGAGATTTGGCA 646
Db 479 ATGGCATTTGCTCGATCTGATCTTCACTTTTAAAGAGCTTCTGACTGCTGTTTTCTCCA 538
Qy 647 AGGTAGCCAGAAATTCATTCGGAATAGCCAAATATGTCCTTCTTCATCTTCAACACC 706
Db 539 AAGTATCTCTCCGACAAATAGATCTGTTCTGTAAAGTCTCTGCAATTTACATGATC 598
Qy 707 TGGATCTGGGTTTTACCTAGACAGACAGCGGAGCTTTATCTAAGGCTATTGACAGAG 766
Db 599 TTGATCTTCTGTTACATTTGAATCGAAGAACAGAGAGCTTGAATAGATCATGACCGTG 658
Qy 767 GAACAAGGGGTATCAGTTTGTCTGAGTGTGATTTAAATCTTCTCCATCATGT 826
Db 659 GAAGTGGCGGATTAACAGATCTCTCTGTAATGTTTCAATGTTTCCAAACATCT 718
Qy 827 TTGAAGTATGCTGTTCAGTGTGTTTGTATTAACAATGCGGTGCCAGTTGCTTTGG 886
Db 719 TGGAGATATCAATGATGACAGATATATGACATACAAATTTTGAACCGTTTGGCTTTGA 778
Qy 887 TAACCTTGAACAATTTGATATACACAGCATTCACAGTTCACACAGCGTGGAGAA 946
Db 779 TCACTTCTCTTCCGTTGATATCATATACATCTTTCACATTTGTGTGACACAGATATGAA 838
Qy 947 CTAGATTGAATAGAAATGAACAAGACAGATATGATGACAGGAATAGCTGCTATAGACT 1006
Db 839 CTAAAGTTTAAAGACGTATACCAAGCTGATATGATGACACAAAGAGCTATTGATT 898
Qy 1007 CACTGCTGAATATGAACCTGGAAGTATTTTAAATATGAAGAATATGAAGACAGAGAT 1066

```

Db	889	CTCTGTAAACTACGAGACTGTTAAGTATTTTAAACAGAAAGTTACGAAGCGAAGACT	958
Qy	1067	ATGATGATTTTTTGAAGACGTATGAGACTGCTCATTTGAAAAGTACTCTACTCTGGCTA	1128
Db	959	ACGATGATTTACTCGAAGATATGAGATGCTGCATTGGCAAATCAGAAAGACTTGGCT	1018
Qy	1127	TGCTGAACCTTTGGTCAAAAGTCTATTTTCAGTGTGGTTTTACGCTATATATGTGTCTG	1188
Db	1019	TTTTAGATTTTGGACAGAGTTTTATATTTAGTACAGCTTTGTCTACATCAATGAGTGGTGT	1078
Qy	1187	CCAGTCAGGGAATTTGGCAGGTACCTCTTACTGTGGAGATCTAGTAATGTGAATGAC	1248
Db	1079	GTTCTCAAGGATTTATGAATGAGAGATGACAGTTGGTGAATTTGGTGAATGTGCT	1138
Qy	1247	TGCTTTTTCAGCTTTTCATTTACCCCTGAACTTTCTGGGAACGTATATATGAGACTAGAC	1308
Db	1139	TTCTCTTTCAATGTGCTCTTCTCTTACTTTCTCGGTGGTGTTTACGTGAACCGTTCT	1198
Qy	1307	AAGCATCATATGATATGATGACACTTGTTTACTCTACTCAAGGTATGACATCCCAATTAAG	1368
Db	1199	AGGGCTCTTGTGATATGATAAATCATTTGTTCAGTGTCTAGAGGAGAGATCCGATATTTGAG	1258
Qy	1387	ACAAAGTATGAGCATCTCCCTTCAGATTCACACACAGACGCT---ACCGAGGCTTTG	1422
Db	1289	ATPAAAGATACAGACACAAGCTACCTCTCTTGTCTGAGAGGTGGAACATCAGTTTG	1318
Qy	1424	ATATATGCTATTTTGAATATACATTTGAGGGCAGAAAGTCTTATGTGAATATCTTTTGAAG	1483
Db	1319	AGAAATGATCATTTTATAGTTATCTGCAAGAGAAAGATTTTGAATGGGATTTCAATTTGAG	1378
Qy	1484	TCCCTGCAGAAAGAAAGTGCCACTTGTAGAGAGTATGGGTACAGGAAAGACAAATAG	1543
Db	1379	TACCCGGCAGGAAAAAGTGATGCGATTTGTGTATGCGGAACGGCAATTCAGACATTC	1438
Qy	1544	TGAGGCTATATTTTCGCTTCATAGGCTCAAAAAGGTATACATTTATCTGTGCTGCAAA	1603
Db	1439	TTAGATATATTTTCAGGTTCTTTTACACAGATTTCTGTAATGTTAGCATTTAGTGCAG	1498
Qy	1604	ATATACAGATGTGAGCCTCGAAGAAAGCTTGCAGAGGAGTGGAGTGTACTCAGAGATG	1663
Db	1499	ATATPAAAGAGTACACTAGAGAGTCTTGGAAGTGTGATGGAGTGTTCCTCAAGATA	1558
Qy	1664	CTGTCTCTTCCATATATCTATTTATTTATTAACAACCTCTTATATGGAACATCAGTCTTAC	1723
Db	1559	CTGTGCTCTTCAATGACACAATATTTTACAACAATTCACATATGGAACCTTTACAGAACAG	1618
Qy	1724	CTGAGGAGGTATGACAGTGGCAAAATTTACTGTGACTTCAATGATGAATTTCTTGCAGTGC	1783
Db	1619	AAGAAGAGGTATGACATGACACTATGAGGACAGTATTCACGATATGATATGAATTTTC	1678
Qy	1784	CACATGATATGACACCCCAATGAGGGAAGAGGAACTCAAGCTTTTACAGAGAGAAAAAGC	1843
Db	1679	CGGATTAATTTTCGACCCCGGTGGGAAGAAAGGGCTGATGCTGATGTGTGGAGAAAAAC	1733
Qy	1844	AAAAGTAGCAATTTGCAAGACCATTTTGAAGAACCCCAAGTCATATCTTATGATGAAG	1903
Db	1739	AAAAGTGGCACTTGTGCTGTCTCTTTCTGAATTCACAGCGAATTTCTTGTGTGAAGAG	1798
Qy	1904	CTACTTCATCGTATGATTCGATTACTGAAGAGACTATTTTGTGTGCTAGGAAGATGTGG	1968
Db	1799	CAACGAATGCCCTGTATATGTAAACGAGGACAGATATTAAGAAAGTTTCAGGTGCGCTAG	1858
Qy	1964	TCAAAACAGAACTTCTATTTTTCATGACACACAGATTTGTCAACAGTGTGATGACAGATG	2022
Db	1859	CGAGTAAAGAACATGCAATTTTCACTTCAACATAGGCTGACATCGATGACATGTGATG	1918
Qy	2024	AAATCATTTGCTTGGATCAGGGTAAAGTAGCCGAACGTGTACCTCCACATGTGTTGCTTG	2083
Db	1919	AGATATATGTAAATGAGAAAGGAAAGGTGGTGAAGAAAAAGGACACACCAAGTGTGTTGG	1978
Qy	2084	CTAACCTCATAGTATCTATTCAGAAATGTGGCATTCACAGACAGCCGTGTG	2136

Db	1979	AGAAATCTGGCAGATACGCTAAATTATGAGACAAACAAAGATACACTGGAG	2031
RESULT 13			
LOCUS	ATH272202	2452 bp	MRNA linear PLN 06-MAR-2001
DEFINITION	Arabidopsis thaliana mRNA for mitochondrial half-ABC transporter (STAI gene).		
ACCESSION	AU272202		
VERSION	AU272202.1	GI:9187882	
KEYWORDS	mitochondrial half-ABC transporter; STAI gene.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 2452) Kushnir S., Babychuk E., Storchenko S., Davey M.W., Penedrobrock J., De Rycke R., Engler G., Stephan U.W., Kispal G., Lill R. and Van Montagu M.		
TITLE	A mutation of the Mitochondrial ABC Transporter Stai Leads to Dwarfism and Chlorosis in the Arabidopsis Mutant stak1		
JOURNAL	Plant Cell 13 (1), 89-100 (2001)		
PIUMED	1158531		
REFERENCE	2 (bases 1 to 2452) Kushnir S.G.		
AUTHORS	Direct Submission		
TITLE	Submitted (14-FEB-2000) Kushnir S.G., Plant Genetics, VIB,		
JOURNAL	University of Gent, Ledeganckestraat 35, B9000, BELGIUM		
FEATURES	Location/Qualifiers		
SOURCE	1..2452		
gene	/organism="Arabidopsis thaliana"		
CDS	/db_xref="taxon:3702"		
	27..2213		
	/gene="STAI"		
	27..2213		
	/gene="STAI1"		
	/function="Fe/S cluster biogenesis"		
	/codon_start=1		
	/evidence=experimental		
	/product="mitochondrial half-ABC transporter"		
	/protein_id="CA87048.1"		
	/db_xref="GI:9187883"		
	/db_xref="SPTREMBL:Q9L788"		
	/translation="MSRSGRFVAPAGLLCRVNLQPPKIPSFYSLSRSDYRLNGFS		
	NYRNSIRTSPIVAINFLSDNSPSPSPPIFFVORSSMLNRLSTSPNDQTT		
	KTEIKITSSDSAMADMKILRTLAGYLVMRNPREFRVIALGLVGAQVLAIVQV		
	PLPLKAVDLMLASATGASLITPLAIPNPLIVPATPAVALIVGIGARTGSSAFNL		
	RTAVFSKVALARTIRSVSRKVFSLHLDLRLYHLSRENGINRIIDGSRAPNIFLSM		
	VFNVPVLTLEISWGSGLIAFGAARPAWITSLSVSTIVTTLAVTORRTFSKANKA		
	DNDSTRAIDSLINVERVKKYFNNGEAEKYDQFLKKYEDPALQVTKSLFLNGOSI		
	I FSTALSTAMVLCSQIIMGQMTVGLVMVNGLLPOLSLNFGSVYRETIQSLVDM		
	KSMFOLLEEKSDITNTSDAKPLVKGGINIEFVNFPEYLPERKILDGI SFVPAGKSV		
	AIYGTCSGSKSTILMRFPEDPDSDIITIDGDIKEVRLDSGSSIGVVPDTPVFN		
	DTIFPHNHYGRILSATEBEVYEARRAIIHTTIDGDIKEVRLDSGSSIGVVPDTPVFN		
	ALARTFLKSPAILLCDEATLSDSTTEALILNLMLKAASRTSI FIAHRLTIAVQCDDE		
	IIVLENKGVVEQGHDELLKSGSRVQLWTQONSVDMLDAAKLE"		
BASE COUNT	703 a 519 c 537 g 693 t		
ORIGIN			
Query Match	23.0%; Score 540.4; DB 8; Length 2452;		
Best local similarity	57.4%; Pred. No. 9.5e-122;		
Matches 1067; Conservative	0; Mismatches 766; Indels 27; Gaps 4;		
307	GGATCCACACAGACCCAAAGAGGTTAATAAGATGTTACTCGGAATATCATATAAA	366	
Db	342 GAATCAAGACAACTCGTCGGATTCAAGACTCACTATAGGCTGACATGAAATCTCCGT	401	
Qy	367 GCAATGCTTTCTTATNGTGTGGCCCAAGACAGCCAGATCTACAGAGTAGAGTCCATT	426	
Db	402 ACGCTGTGCTGCTATTTGTGATGAGAGACATCCAGAAATTCGTTTACAGGTCATGCT	461	
Qy	427 TCGCTGGAGATTTTGGGTGTGCAAGGCCATGAATATGTGTGTTCCCTTATGTTTAA	486	

Db 462 GCTTGGGTTTCTCGTGGGTGCTAAGGTCATGATGTCAGATCCCTTCTCTTCAA 521
Qy 487 TATGCTGAGACAGCT-----CAACGAGTGGGGAACAATGTAACCGAGTG 538
Db 532 TTAGCTGTGATGCTTGTGCTTCACTTCCGTTAGAGCTCTTACCACTTGGCT 581
Qy 539 ATG-----CACCAATATACAGTTGCAACATGAGCAAGAGTCTGATGGCTAT 588
Db 582 GCTACAAACCAACCTCTTACTGTTTTTGTGACTCTGCGGCTGTTCTCATGGATAC 641
Qy 589 GGTGATTAAGAGCTGAGAGCTGTTTTTTTAAACGAAGTTCGAATGCAATTTGGCAG 648
Db 642 GGGATGCTGGACGGGTTCTTGTCTTTATGAACTTGTACTGCTGTTCTCTAAG 701
Qy 649 GTACCCCAAGATTCATCCGAAGATAGCAAAATGCTTCTTCATCTTCAACACCTG 708
Db 702 GTACTCTCCGAACAAATACATCTGTTCTTGAAGAGTGTCTCACTTTTCAATGATTT 761
Qy 709 GATCTGGGTTTCACTGAGCAGACAGCGGAGCTTATCTAAGGCTATTGACAGAGA 768
Db 762 GACCTTGCTTACCACTTATGCGAAGAACAGCGGCTTGAACAGGATTAATGATCGGT 821
Qy 769 ACAAGGGTATCACTTTTGTCTGAGCTTGGTATTAACTTTTCCATCATGTTT 828
Db 822 AGCGGTGCAATTAATTTATCTCTCACTATGTTTCAATGTTGTTCTTACATATTTG 881
Qy 829 GAAGTATGCTTGTCAAGTGTGTTTGTATTAACAATGCGGTGCCAGTTGGTGTGTA 888
Db 882 GAGATATTTATGATATCAGATATCTGCAATACAGATTGAGCTGCTTGGCTGAGTT 941
Qy 889 ACCCTTGGAACCTTGTATACATACACAGATTCACAGTTGAGTCAACGCTGGAAC 948
Db 942 ACTTCTCTCTCGTTGATATATATATGTTTCACTTATGCTGACACATGCGGAAC 1001
Qy 949 AGATTATGAATGAAGTGAACAAAGCAGATATGATGATGATGATGATGATGATGAT 1008
Db 1002 AAGTTATGATGAGCTATGAACAAAGCTGATGATGATGATGATGATGATGATGAT 1061
Qy 1009 CTGCGAATTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Db 1062 CTGATTAATTTATGAGCTATGATTAATTTTAAATGAGGTTATGAGGCTGAAGAT 1121
Qy 1069 GATGATTTTGAAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
Db 1122 GACAGATTCCTGAAGAAATATGAGATGCTGCTGCAATCTCAAAGAGCTTGTTC 1181
Qy 1129 CTGAATTTGCTCAAGTGTATTTTCACTGCTGCTTAACTGATTAATGATGATGAT 1188
Db 1182 TTAAATCTTGAACAAAGTATCATATTTATGATGATGATGATGATGATGATGAT 1241
Qy 1189 AGTGAAGGAATGAGAGGAGTACCTTATGATGATGATGATGATGATGATGATGATGAT 1248
Db 1242 TCTAGGGTATCATGATATGCGCAGATGACATGCTGATGATGATGATGATGATGAT 1301
Qy 1249 CTTTTCAGCTTCAATTAACCTGATCTTGGGAATGATATGATGATGATGATGATGAT 1308
Db 1302 CTTCTTGAAGCTGCTCTCTTAAATTTTGGTATGATGATGATGATGATGATGATGAT 1361
Qy 1309 GCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
Db 1362 AGCCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
Qy 1369 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1428
Db 1422 ACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
Qy 1429 GTGATTTTGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
Db 1476 GTACATTTCACTTCTTCCAGAGAAATATTAATGATGATGATGATGATGATGATGAT 1555
Qy 1489 GCAGAAAGAAAGTGGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548

Db 1536 GCAGAAAGAAAGTGGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1595
Qy 1549 CTATATTTTGGCTCTATGAGCCCTCAAGAGGATGATGATGATGATGATGATGATGATGAT 1608
Db 1596 ATGCTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
Qy 1609 CAAGATGAGCCCTGGAAGCTTGGAGGCACTGGAGTGTACTCAGATGCTGTC 1668
Db 1656 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1715
Qy 1669 CTTTCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728
Db 1716 CTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
Qy 1729 GAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
Db 1776 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
Qy 1789 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
Db 1836 AAATATTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1895
Qy 1849 GTAGCATTTGCAAGAGCATTGTAAGAGACCCCACTCATCTTATGATGATGATGATGAT 1908
Db 1896 GTGCACTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1955
Qy 1909 TCATGTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1968
Db 1956 AGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2015
Qy 1969 CACAGAACTTCTATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
Db 2016 AACAGCATCTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2075
Qy 2029 ATTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
Db 2076 GTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2135
Qy 2089 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2148
Db 2136 TC--AGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2192

RESULT 14
AF287698
LOCUS 2043 bp mRNA linear PLN 02-SEP-2000
DEFINITION Arabidopsis thaliana half-molecule ABC transporter ATM2 mRNA,
complete cds.
ACCESSION AF287698
VERSION AF287698.1 GI:9964118
KEYWORDS
ORGANISM Arabidopsis thaliana.
SOURCE Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2043)
AUTHORS Sanchez-Fernandez,R., Mari,S., Dancis,A. and Rea,P.A.
TITLE Functional half-molecule ABC transporters from Arabidopsis: the ATM
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2043)
AUTHORS Sanchez-Fernandez,R., Mari,S., Dancis,A. and Rea,P.A.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2000) Plant Science Institute, Department of
Biology, University of Pennsylvania, 3800 Hamilton Walk,
Philadelphia, PA 19104-6018, USA
FEATURES
source
1..2043
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/note="ecotype: Col-0"
CDS
1..2043

VERSION AC018231.1 GI:6552960
 KEYWORDS HTG; HTGS_PHASE2.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 50701)
 AUTHORS Adams,M. and Venter,J.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 COMMENT This sequence was identified as CD1:10213948 by the submitter.
 For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source 1. .50701
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 13744 a 12081 c 12135 g 12741 t
 ORIGIN
 Query Match 20.5%; Score 480.6; DB 2; Length 50701;
 Best Local Similarity 59.6%; Pred. No. 5.2e-107;
 Matches 810; Conservative 0; Mismatches 549; Indels 0; Gaps 0;
 QY 587 ATGTGATCAAGAGCTGGAGCTGTTTAAAGAAAGTTGCAAAATGCAAGTATTTGGCA 646
 DB 11810 ATGGCATGCTAGAGCCGAGTGGGGCTTTAAATGAAGTCCGAAATGCAAGTGGCA 11751
 QY 647 AGTAGCCCAAGATTCATCCGAAAGATAGCCAAATATGTTCTTCATGTTGCAACC 706
 DB 11750 AGGTGGCCCACTCTGATCCGAAAGATCGCCAGCAATGTTCTTGCATGCAACC 11691
 QY 707 TGGATCTGGGTTTCACTGAGCAGACGAGCGAGCTTTATCTAAGCTATTTGACAG 766
 DB 11690 TGGATCTGGGCTTCACTGAGCAACAACTGAGGAGCTGTCAAAAGCAATGATGAG 11631
 QY 767 GAACAAAGGGATCAAGTTTGTCTGAGTGTGTTGTTAATCTTCTTCCATCATGT 826
 DB 11630 GATGAGGGGATTAACCTTGTCTCCGCAATGCTTCAACATTTGTGCCCATATCT 11571
 QY 827 TTGAAGTATGCTGTGAGTGTGTTTGTATTAACAATGGGGGCCAAGTTGCTTTGG 886
 DB 11570 TTGAGTGGCCCTGTGTCCAGTATCCGGGAGTGAAGTGTGGCTTGTGCTGGTG 11511
 QY 887 TAACTTGGAACTTGTATCATATACAGCATTCACAGTTGACATCAACGTTGAGAA 946
 DB 11510 TCACGATGGGCTGGTGGCATATACGCTGCTACCTGAGTGTACCAAGTGGGCA 11451
 QY 947 CTAGATTTAGATAGAAATGAACAAAGCATTAATGATGAGTAAATGCTTATAGACT 1006
 DB 11450 CCCAGTTCCGATTTATTAAGAACAGGGGAGAAAGAGCGGCAACAGGCCGTAGACT 11391
 QY 1007 CACTGCTGAATTAAGAACTGAAATTTTAATTAATGAAGATAGAAACACAGAGAT 1066
 DB 11390 CGCTGATCAACTACGAGCGGTGAAGTACTTCAACAAAGAAAGTACGAGCGGGTGTCT 11331
 QY 1067 ATGATGATTTTGAAGCGATAGACTGCTTATTTGAAGAAAGTACCTTACTCTGCTA 1126
 DB 11330 ACAACGAGGTCTGAAAGATATAGAGGCGCACACTTGAAGACTAGCTCCAGTCTGCTC 11271
 QY 1127 TGTGAACCTTTGTCAAAGTCTATTTTCAAGTGTGGTTTAAAGCTAATAATGTTGCTCG 1186
 DB 11270 TGTCAACTTTGGCCAAAGCCATTTTCAAGCAGCGCACTGATCTGATTAATGTAATGG 11211
 QY 1187 CCACTGAGGAATGTGGCAGGTACCTTACTGTTGAGATCTAGTAATGTTGATGAC 1246
 DB 11210 CCGCAAGAGATTTGCCAAGGCAACATGACGTTGGATTTGTGATGTTCAAGCCC 11151

QY 1247 TGTCTTTCAAGCTTTCATTAACCCCTGAACCTTTCTGGAACTGTATATAGAGACTAGAC 1306
 DB 11150 TGTCTTTCAAGCTTTCATTAACCCCTGAACCTTTCTGGAACTGTATATAGAGACTAGAC 11091
 QY 1307 AAGCACTCATAGATATGAACACCTGTTTACTACTCAAGGTAGACCAACCAATTAAG 1366
 DB 11090 AAGCTTTGCTGACATCGGGCCATGTTTCACTGTTGAAAGCTGACAGTACATTGCA 11031
 QY 1367 ACAAGTATGATGATCTCCCTTCAGATCAACCAAGACAGTACCGTGGCTTTGATA 1426
 DB 11030 CAGCCGCAATGCCCACACCTCGTGTGAGACACACTAATCTATCATTAAGTTCGCCA 10971
 QY 1427 ATGTGCAATTTTGAATATGATGAGGCGCAAAAGTCTTATGTAATTCCTTGAAGTCC 1486
 DB 10970 ACGTGAACCTTCGATGACAACTGGCAAGCCATTTTCCGGGACCTCAGTTTCACTATAC 10911
 QY 1487 CTGAGGAAGAAAGATGAGCCATGTTGAGAGTATGAGGTGAGGAAAGCAATATGTA 1546
 DB 10910 CCGCCGGTAAAGCTGCAATGTTGGGGGCTGTGGCTCAGGAAATTCGTGATGGTGC 10851
 QY 1547 GGTATTAATTTGCTCTATGAGGCTCAAAAGGATAGCATTTATCTTGTGTCATAATA 1606
 DB 10850 GCTTCTCTCCGCTTTTGAAGCAACCTGTGTAAGTGTGATGGTGGCCAGATA 10791
 QY 1607 TACAAGATGATGAGCTGGAAGACCTTTCGAGGCGAGTGGAGTGTACTTCAGAGTCTG 1666
 DB 10790 TCAGCGCGGTGATTTGAGAGCCTGTGGCAAGGTATTTGACAGTGTGCCGAGATTCGG 10731
 QY 1667 TCTCTTCCATTAATATATTTATTAACCTCTTATATGAAACATCAGTGTTCACCTG 1726
 DB 10730 TGTATTTCCAAACACCATGAGACAAATATCACTATGGAACCTGTCCAAATGCCAGC 10671
 QY 1727 AGGAAGTATGATGAGTGGCAAAATTAAGCTGATGATGATGCAATCTTTCGAATGCCAC 1786
 DB 10670 CAAAGTGCAGAAATGCTGCCCGCATGCTGATTTGATGATGATGATGATGATGATGATG 10611
 QY 1787 ATGATATGACACCCCAAGTGGGAAAGAGAGTCAAGCTTTCAGAGAGAAAGCAAA 1846
 DB 10610 GACAGTACTCAGTCAAGTGGGCAAGTGGATTAAGTTATCCGGTGGCGAGAAAGCAGC 10551
 QY 1847 GAGTAGCAATTCGAAGGCAATTTTGAAGACCCCGCAATCTTATGATGAACTA 1966
 DB 10550 GCGTTGCATCGCTAGGGCCATCTGAAAGAACACCCCATTTCTGATCTTTGAAGAGCTA 10491
 QY 1907 CTTCATGTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1945
 DB 10490 CAAGCACTGAGCTGATCAATCAGAGCATGTGATGTTG 10452

Search completed: February 15, 2003, 03:21:32
 Job time : 6092 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 00:34:00 ; Search time 480 Seconds
(without alignments)
11001.950 Million cell updates/sec

Title: AF133659
Perfect score: 2345
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAATCATCAATT 2345

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001.DAT:*
23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2277.6	97.1	2384	20	AAK36895	Human ABC-Transporter
2	2248	95.9	2407	20	AAV82520	Human ATP-binding
3	2248	95.9	2407	21	AAK62411	Human ATP-binding
4	2150.4	91.7	2193	22	AAH14279	Human CDNA sequence
5	564.8	24.1	2182	23	ABL01991	Drosophila melanog
6	564.8	24.1	2539	23	ABL01989	Drosophila melanog
7	551.6	23.5	620	22	AAH05966	Human CDNA clone
8	517	22.0	807	22	AAH34971	Human colon cancer
9	480.6	20.5	6534	23	ABL01988	Drosophila melanog

10	480.6	20.5	6541	23	ABL01990	Drosophila melanog
11	408.6	17.4	437	21	AAA45541	Human secreted exp
12	358.6	15.3	5369	20	AAK70880	A. gossypii ADEA D
13	316.6	13.5	1185	21	AAK37701	Arabidopsis thaliana
14	299.6	12.8	2770	23	ABL28787	Drosophila melanog
15	284.2	12.1	2984	24	AAK27254	Human transporter
16	282.6	12.1	1725	21	AAZ94743	Human ATP binding
17	282.6	12.1	1933	21	AAZ48823	CDNA encoding a hu
18	282.6	12.1	1958	22	AAI58469	Human ABC transporter
19	282.6	12.1	2529	22	AAI51604	Human secreted pro
20	277.6	11.8	296	20	AAK41426	Human ABCB12 trans
21	274.8	11.5	2893	22	AAK00010	Drosophila melanog
22	269.2	11.5	5383	23	AAI28786	Human polyomavirus
23	255.6	10.9	1960	22	AAI60255	S. epidermidis ope
24	189.2	8.1	1647	22	AAH52466	S. epidermidis gen
25	189.2	8.1	3039	22	AAH34851	S. epidermidis gen
26	189.2	8.1	3585	22	AAH54600	S. epidermidis gen
27	189.2	8.1	3736	22	AAH54207	Staphylococcus epi
28	188.8	8.1	1749	24	ABN92430	Staphylococcus aur
29	187.4	8.0	1053	18	AAV75107	Bacillus clausii g
30	185.8	7.9	939	24	ABK79220	Genomic sequence o
31	185.4	7.9	2365589	24	ABA90521	S. aureus polypept
32	184.6	7.9	1737	20	AAK31863	Staphylococcus aur
33	184.6	7.9	1737	23	AAK54766	Enterococcus faeca
34	182.4	7.8	1716	23	AAK52885	Enterococcus faeca
35	180.8	7.7	1716	22	ABA82943	Streptococcus pneu
36	180.2	7.7	7760	19	AAV52196	Listeria monocytog
37	178.6	7.6	1424	24	ABQ70460	ApixB gene. Acti
38	177.8	7.6	2586	24	ABQ70666	A. pleuropneumonia
39	170	7.2	5120	18	AAI73218	Streptococcus pneu
40	170	7.2	5120	19	AAZ68585	E. coli O96 pathog
41	165.8	7.1	13188	19	AAV52203	Listeria monocytog
42	165.4	7.1	3575	19	AAV12855	Bacillus lichenifo
43	158.4	6.8	2832	24	ABQ70839	Enterococcus faeca
44	158.2	6.7	1233	24	ABK74885	
45	158	6.7	1719	22	ABA82948	

ALIGNMENTS

RESULT 1
AAK36895 standard; cDNA; 2384 BP.

AAK36895:

16-JUL-1999 (first entry)

Human ABC-Transporter-7 (HABCT) coding sequence.

HABCT; human; ABC-transporter-7; diagnosis; cancer; autoimmune disease; Addison's disease; insulin-dependent diabetes mellitus; therapy; microosomal disorder; ss.

Homo sapiens.

WO9921885-A1.

06-MAY-1999.

29-OCT-1997; 97WO-CN00120.

29-OCT-1997; 97WO-CN00120.

(UYSH-) UNIV SHANGHAI SECOND MEDICAL.

Zhang Q;

WPI: 1999-312945/26.

F-PSDB; AA14068.

HABCT a protein useful in the treatment of cancer, diabetes,

antibodies

1022

microsomal disorders and Addison's disease
Claim 2; Page 26-27; 31pp; English.

This sequence encodes the human ABC-transporter-7 (HABCT) protein of the invention. HABCT is an ABC transporter protein. The DNA, vectors containing it and recombinant host cells are useful for recombinant production of HABCT. The DNA, HABCT and antibodies against HABCT are useful as research reagents, for screening assays and in diagnostic assays. Antagonists and agonists of HABCT can be used to inhibit or enhance, respectively, the activity of HABCT or its fragments can be coding sequence. Anti-HABCT antibodies and HABCT or its fragments can be used in vaccines. In particular, the proteins, antibodies, agonists and antagonists can be used for treating, e.g. cancer, autoimmune disease, Addison's disease, microsomal disorders and insulin-dependent diabetes mellitus, related to both an excess and insufficient amounts of HABCT.

Sequence 2384 BP; 719 A; 467 C; 567 G; 631 T; 0 other;

Query Match 97.1%; Score 2277.6; DB 20; Length 2384;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2318; Conservative 0; Mismatches 24; Indels 6; Gaps 2;

```
OY 1 ATGGGCTGCTCGCGATCATCTTGGCGCTGGCGCGCGCGCTCTTTCGAAAAG 60
DB 9 ATGGGCTGCTCGCGATCATCTTGGCGCTGGCGCGCGCGCTCTTTCGAAAAG 68
OY 61 CGCGGCACTCGCGATTCTGATCCGCCCTTTAGTCTCTTTAGCGCTCAGGTCCGAG 120
DB 69 CGCGGCACTCGCGATTCTGATCCGCCCTTTAGTCTCTTTAGCGCTCAGGTCCGAG 128
OY 121 TGGAGGCCATCACTCGCGCCTTGGGAACGCTCGAGCTACAGATTCAGAGTCA 180
DB 129 TGGAGGCCATCACTCGCGCCTTGGGAACGCTCGAGCTACAGATTCAGAGTCA 188
OY 181 TTAATAAGATCAGATGCGAGATTGGGAAAGCAATTCAGACAAGTCTTATGATGT 240
DB 189 TTAATAAGATCAGATGCGAGATTGGGAAAGCAATTCAGACAAGTCTTATGATGT 248
OY 241 GCAAAAGCTCTCCAGGTATGCGCACTGATGAAAAAGAGACATGTTGGATGTCATGA 300
DB 249 GCAAAAGCTCTCCAGGTATGCGCACTGATGAAAAAGAGACATGTTGGATGTCATGA 308
OY 301 GGAGAGGACTCCACACACACCAAAAGAGGTTAAAGATGTTGATCTCGAAAAATC 360
DB 309 GGAGAGGACTCCACACACACCAAAAGAGGTTAAAGATGTTGATCTCGAAAAATC 368
OY 361 ATAAAAGCAATGCTTTCTATGTTGGCCCAAAAGACAGGCCAGATCTACAGACTAGATT 420
DB 369 ATAAAAGCAATGCTTTCTATGTTGGCCCAAAAGACAGGCCAGATCTACAGACTAGATT 428
OY 421 GGCATTTGCTGGGATTTTGGTGGTGCMAAGGCCATGAAATTTGGTTCCTTCATG 480
DB 429 GGCATTTGCTGGGATTTTGGTGGTGCMAAGGCCATGAAATTTGGTTCCTTCATG 488
OY 481 TTTAAATATGCTGTAGACAGCCTCAACAGATGTCGGGAAACATGCTGAACCTGAGTAT 540
DB 489 TTTAAATATGCTGTAGACAGCCTCAACAGATGTCGGGAAACATGCTGAACCTGAGTAT 548
OY 541 GCACCAATATACAGTTGCAACCATGCAACAGCAGATTCTGATTTGGCTATGCTATCAAGA 600
DB 549 GCACCAATATACAGTTGCAACCATGCAACAGCAGATTCTGATTTGGCTATGCTATCAAGA 608
OY 601 GCTGAGCTGCTTTTAAACGAAGTTCAAAATGCTATTTGGCAAGTACCCGAGAT 660
DB 609 GCTGAGCTGCTTTTAAACGAAGTTCAAAATGCTATTTGGCAAGTACCCGAGAT 668
OY 661 TCAATCCGAAGATAGCAAAATGCTTTCTCATCTTCAACAACCTGGATCTGGGTTTT 720
DB 669 TCAATCCGAAGATAGCAAAATGCTTTCTCATCTTCAACAACCTGGATCTGGGTTTT 728
OY 721 CACCTGAGCAGACAGCGGAGCTTTATCTTAAGCTATTGACAGAGAAACAAAGGGTATTC 780
```

```
DB 729 CACCTGAGCAGACAGCGGAGCTTTATCTTAAGCTATTGACAGAGAAACAAAGGGTATTC 788
OY 781 AGTTTGTCCCTGAGTCTTGGATTTAACTCTTCCCATCATGTTTGAAGTATGATG--- 837
DB 789 AGTTTGTCCCTGAGTCTTGGATTTAACTCTTCCCATCATGTTTGAAGTATGATGCTT 848
OY 838 CTGTGAGTGTGTTTGTATTAACAAATGGGTGCCAGTTTGGCTTTGATACCTTTGGA 897
DB 849 CTGTGAGTGTGTTTGTATTAACAAATGGGTGCCAGTTTGGCTTTGATACCTTTGGA 905
OY 898 ACACCTTGTATCAACACAGATTCACAGTTGCAAGTTCACACGGTGGAGAACTAGATTAGA 957
DB 906 ACACCTTGTATCAACACAGATTCACAGTTGCAAGTTCACACGGTGGAGAACTAGATTAGA 965
OY 958 ATGAATAAGCAACAGAGATTAATGATGCAAGTATGCTGCTATGACTCAGTCTGAAT 1017
DB 966 TTGGAATTTGACCAAGCAGATTAATGATGCAAGTATGCTGCTATGACTCAGTCTGAAT 1025
OY 1018 TATGAAACTGTGAAGTATTTTAAATATGAAGAATATGAAACACAGATATGATGATTT 1077
DB 1026 TATGAAACTGTGAAGTATTTTAAATATGAAGAATATGAAACACAGATATGATGATTC 1085
OY 1078 TTGAAAGCTATGAGACTGCTTCAATGAAGAAGTACCTTACCTCGGCTATGCTGAATTT 1137
DB 1086 TTGAAAGCTATGAGACTGCTTCAATGAAGAAGTACCTTACCTCGGCTATGCTGAATTT 1145
OY 1138 GGTCAAGTGTATTTTCACTGTGCTTTAAACAGTATTAATGATGCTCGCAGTCAAGGA 1197
DB 1146 GGTCAAGTGTATTTTCACTGTGCTTTAAACAGTATTAATGATGCTCGCAGTCAAGGA 1205
OY 1198 ATTGTGCAAGTACCTTACTGTTGAGATCTAGTAATGTTGAATGGAATGCTTTTTCAG 1257
DB 1206 ATTGTGCAAGTACCTTACTGTTGAGATCTAGTAATGTTGAATGGAATGCTTTTTCAG 1265
OY 1258 CTTTCAATTAACCCCGAATCTTTCGGGAACCTGTATATGAGAGACTAGACAAGCACTATA 1317
DB 1266 CTTTCAATTAACCCCGAATCTTTCGGGAACCTGTATATGAGAGACTAGACAAGCACTATA 1325
OY 1318 GATATGAACACCTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1377
DB 1326 GATATGAACACCTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1385
OY 1378 GCATCTCCCTTCAAGATCACACACAGACAGTACCGTGGCCTTTGATATATGTCATTTT 1437
DB 1386 GCATCTCCCTTCAAGATCACACACAGACAGTACCGTGGCCTTTGATATATGTCATTTT 1445
OY 1438 GAATACATTTGAGGGCCAGAAAGTCCCTTATGGAATATCCTTTGAAGTCCCTGACAGAAAG 1497
DB 1446 GAATACATTTGAGGGCCAGAAAGTCCCTTATGGAATATCCTTTGAAGTCCCTGACAGAAAG 1505
OY 1498 AAAGTGGCATTTGTAGAGATGAGTGGGTCAAGGAAAGACAAATAGTGAAGCTATTTAT 1557
DB 1506 AAAGTGGCATTTGTAGAGATGAGTGGGTCAAGGAAAGACAAATAGTGAAGCTATTTAT 1565
OY 1558 CGCTTTCTATGAGCTCAAAAAGGTAGCATTTATCTGTGGTCAAAATATATCAAGAGTGTG 1617
DB 1566 CGCTTTCTATGAGCTCAAAAAGGTAGCATTTATCTGTGGTCAAAATATATCAAGAGTGTG 1625
OY 1618 AGCCTGAAAGCCTTCGAGGGGCAAGTGGAGTGTGATCCTCAGATGCTGTCTCTTCAT 1677
DB 1626 AGCCTGAAAGCCTTCGAGGGGCAAGTGGAGTGTGATCCTCAGATGCTGTCTCTTCAT 1685
OY 1678 AATACCTATTTATCAACCTCTTATATGAAACATCAGTGTCTTACCTGAGAGAGTAT 1737
DB 1686 AATACCTATTTATCAACCTCTTATATGAAACATCAGTGTCTTACCTGAGAGAGTAT 1745
OY 1738 GCAAGTGGCAAAATTAAGTGTGACTTATATGCAATTTCTTGAATGCCACATGAGATATAC 1797
DB 1746 GCAAGTGGCAAAATTAAGTGTGACTTATATGCAATTTCTTGAATGCCACATGAGATATAC 1805
OY 1798 ACCCAAGTAGGGAACGAGAGCTCAAGCTTTCAGGAGGAGAAACCAAGAGTACAAAT 1857
DB 1806 ACCCAAGTAGGGAACGAGAGCTCAAGCTTTCAGGAGGAGAAACCAAGAGTACAAAT 1865
```

QY 1858 GCAAGAGCCATTGTAAGAGACCCCCAGTCACTATGATGTAAGTACTTCACTGTTA 1917
1866 GCAAGAGCCATTGTAAGAGACCCCCAGTCACTATGATGTAAGTACTTCACTGTTA 1925
QY 1918 GATTGATGTAAGAGACCTATCTGGTCCATGTAAGTGTGTCAACACAGAACT 1977
1926 GATTGATGTAAGAGACCTATCTGGTCCATGTAAGTGTGTCAACACAGAACT 1985
QY 1978 TCTATTTTCATTGACACAGATGTCACAGTGGTGTGATGACATGTAAGTACTTGTCTTG 2037
1986 TCTATTTTCATTGACACAGATGTCACAGTGGTGTGATGACATGTAAGTACTTGTCTTG 2045
QY 2038 GATGAGGTAAAGTAAAGCCGGAAGCTGTGTAACCAACATGTTGCTTGTCTTCACTAGT 2097
2046 GATGAGGTAAAGTAAAGCCGGAAGCTGTGTAACCAACATGTTGCTTGTCTTCACTAGT 2105
QY 2098 ATCTATTCAGAAATGTGTCATACACAGAGCCGCTGTGTCGAACCATGATTAACCCCAA 2157
2106 ATCTATTCAGAAATGTGTCATACACAGAGCCGCTGTGTCGAACCATGATTAACCCCAA 2165
QY 2158 TGGGAGCAAGAAAGAAATATATATCCAAAGAGAGGAAAGAAAGAACTACAGAGAA 2217
2166 TGGGAGCAAGAAAGAAATATATATCCAAAGAGAGGAAAGAAAGAACTACAGAGAA 2225
QY 2218 ATGTCAATAGTGTGAAGAGCTGTGAAAGCTGTGCTAAGTCACTAAGACATTTTTC 2277
2226 ATGTCAATAGTGTGAAGAGCTGTGAAAGCTGTGCTAAGTCACTAAGACATTTTTC 2285
QY 2278 TTTTGTGTGTGTGTGTGACATCAATTTTGCATGTAAGCAAGATTTTATTAATAAAT 2337
2286 TTTTGTGTGTGTGTGTGACATCAATTTTGCATGTAAGCAAGATTTTATTAATAAAT 2345
QY 2338 CATACATT 2345
2346 CATACATT 2353
Db

RESULT 2
AAV82520
ID AAV82520 standard; cDNA; 2407 BP.
XX
AC AAV82520;
XX
DT 22-MAR-1999 (first entry)
XX
DE Human ATP-binding cassette transport protein encoding cDNA.
XX
KW Human; ATP-binding cassette transport protein; ABCxH; cancer;
neural disorder; Alzheimer's disease; dementia; depression;
Down's syndrome; epilepsy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 9..2252
FT /tag= a
XX
PN US5858719-A.
XX
PD 12-JAN-1999.
XX
PF 17-JUL-1997; 97US-0895522.
XX
PR 17-JUL-1997; 97US-0895522.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC, Hillman UL, Shah P;
XX
DR WPI; 1999-119878/10.
XX
P-PSDB; AAV8585.
XX

PT Isolated ATP-binding cassette transport protein - used to develop
PT products for treating cancers or neuronal disorders, e.g.
PT Alzheimer's disease, dementia, depression, Down's syndrome or
PT epilepsy
XX
PS Claim 4; Fig 1; 39pp; English.
XX
CC The present sequence encodes human ATP-binding cassette transport
CC protein (ABCxH). The ABCxH proteins are involved in disease-related
CC transport processes, particularly in cancers and neuronal disorders.
CC Antagonists of ABCxH can be used to prevent or treat cancers or
CC neuronal disorders e.g. Alzheimer's disease, amnesia, amyotrophic
CC lateral sclerosis, bipolar disorder, cataplexy, cerebral neoplasms,
CC dementia, depression, Down's syndrome, tardive dyskinesia, dystonias,
CC epilepsy, Huntington's disease, multiple sclerosis, neurofibromatosis,
CC Parkinson's disease, paranoid psychoses, schizophrenia or Tourette's
CC disorder.
XX
SQ Sequence 2407 BP; 726 A; 477 C; 570 G; 629 T; 5 other;

Query Match 95.9%; Score 2248; DB 20; Length 2407;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 8 TGCTGGGATGATCTTCTGTGGCGTGGCGGCGCGCGCTGCTTTGAAAGCGCCGCG 67
Db 1 TGCTGGGATGATCTTCTGTGGCGTGGCGGCGCGCGCTGCTTTGAAAGCGCCGCGC 60
QY 68 ACTCGGATTTGATTCGCGGCTTTAGTCTCTGTGTAGCGGCTCAAGTCCCGCATGGAGGC 127
Db 61 ACTCGGATTTGATTCGCGGCTTTAGTCTCTGTGTAGCGGCTCAAGTCCCGCATGGAGGC 120
QY 128 CACATCACTCGCGGCTTTGGGAAAGCGCTGAGCTTCAAGATTTCCAGATCTTAATAA 187
Db 121 CACATCACTCGCGGCTTTGGGAAAGCGCTGAGCTTCAAGATTTCCAGATCTTAATAA 180
QY 188 GTATCAATGGCAGAGATTGGGAAAGGCAATTCAGACAGTTCTTAGTGTGCAAAAG 247
Db 181 GTATCAATGGCAGAGATTGGGAAAGGCAATTCAGACAGTTCTTAGTGTGCAAAAG 240
QY 248 CTCTCCAGGTATGGCCACTGATGAAAGAGGCAATGTGGCATGTGATGCAAGAGAG 307
Db 241 CTCTCCAGGTATGGCCACTGATGAAAGAGGCAATGTGGCATGTGATGCAAGAGAG 300
QY 308 GACTCCACAGACCCCAAAAGAGGTTAAAGATGTGATCTCGGAAATCATTAAG 367
Db 301 GACTCCACAGACCCCAAAAGAGGTTAAAGATGTGATCTCGGAAATCATTAAG 360
QY 368 CAATGCTTTCTTATGTGTGGCCCAAGACAGGCAATCTACAGAGTGTGCAATTT 427
Db 361 CAATGCTTTCTTATGTGTGGCCCAAGACAGGCAATCTACAGAGTGTGCAATTT 420
QY 428 CGCTGGATTTTGGGTGTGCAAGGCAATGATTTGTGTGCTTCCCTTCACTTTAAT 487
Db 421 CGCTGGATTTTGGGTGTGCAAGGCAATGATTTGTGTGCTTCCCTTCACTTTAAT 480
QY 488 ATGCTGTAGACAGCTCAACCAAGATGTGGGAAACATGCTGAACCTGAGTGAACCAA 547
Db 481 ATGCTGTAGACAGCTCAACCAAGATGTGGGAAACATGCTGAACCTGAGTGAACCAA 540
QY 548 ATACAGTTGCAACCATGGAACAGAGTTCTGATTTGGCTATGCTATCAAGAGCTGAG 607
Db 541 ATACAGTTGCAACCATGGAACAGAGTTCTGATTTGGCTATGCTATCAAGAGCTGAG 600
QY 608 CTGCTTTTAAAGCAAGTTGCAATGCAATGTTGGCAAGTGTGCAAGTGTGCAATTC 667
Db 601 CTGCTTTTAAAGCAAGTTGCAATGCAATGTTGGCAAGTGTGCAAGTGTGCAATTC 660
QY 668 GAAGATAGCAAAATGCTTTCTCCATCTTCAAGCTGATCTGGGTTTCACTGA 727
Db 661 GAAGATAGCAAAATGCTTTCTCCATCTTCAAGCTGATCTGGGTTTCACTGA 720
QY 728 GCAGACAGACGGAGCTTTATCTAAGCTATTGACAGAGAAACAAAGGGATATCAGTTTGG 787


```

Db 721 GCAGACAGCGGAGCTTATCTAAGGCTATGACAGAGAAACAGGGTATCACTTTTG 780
Qy 788 TCCCTAGGCTTTGGATTTTATCTTCCATCATGTTGAAGTGAAGCTGTCAGTG 847
Db 781 TCCCTAGGCTTTGGATTTTATCTTCCATCATGTTGAAGTGAAGCTGTCAGTG 840
Qy 848 GTGTTTGTATTAACAATGCGGTGCCAGTTTGTGTTGAACCCCTGGAACTTTGTA 907
Db 841 GTGTTTGTATTAACAATGCGGTGCCAGTTTGTGTTGAACCCCTGGAACTTTGTA 900
Qy 908 CATACACAGCATTCACAGTTGCACTACACGCTGAGAACTTAGATTGAAATAGA 967
Db 901 CATACACAGCATTCACAGTTGCACTACACGCTGAGAACTTAGATTGAAATAGA 960
Qy 968 ACAAGCGATATGATGAGGATATGCTGCTATGACTACAGCTGGAATTAAGAACTG 1027
Db 961 ACAAGCGATATGATGAGGATATGCTGCTATGACTACAGCTGGAATTAAGAACTG 1020
Qy 1028 TGAATATTTTATATGAAAGATATGAAGACAGAGATATGATGATTTTGAAGAGT 1087
Db 1021 TGAATATTTTATATGAAAGATATGAAGACAGAGATATGATGATTTTGAAGAGT 1080
Qy 1088 ATGAGACTGCTTCATTGAAAAGTACCTCTACTCTGCTATGCTGAACCTTGGTCAAAGTG 1147
Db 1081 ATGAGACTGCTTCATTGAAAAGTACCTCTACTCTGCTATGCTGAACCTTGGTCAAAGTG 1140
Qy 1148 CTATTTTCAAGTCCGTTTAAACAGCTATATATGCTGCTGCAAGGAAATTTGGCAG 1207
Db 1141 CTATTTTCAAGTCCGTTTAAACAGCTATATATGCTGCTGCAAGGAAATTTGGCAG 1200
Qy 1208 GTACCTTACTGTTGAGAGATCTAGTATATGATGATGAGCTGCTTTTACGTTTCATTAC 1267
Db 1201 GTACCTTACTGTTGAGAGATCTAGTATATGATGATGAGCTGCTTTTACGTTTCATTAC 1260
Qy 1268 CCTGAACTTTCTGAGAACTGTATATAGAGAGCTAGACAGCACTCATGATATGAACA 1327
Db 1261 CCTGAACTTTCTGAGAACTGTATATAGAGAGCTAGACAGCACTCATGATATGAACA 1320
Qy 1328 CCTGTTTACTGCTACTCAAGGATGACACCCAAATTAAGACAAAGTGAATGCTCCCC 1387
Db 1321 CCTGTTTACTGCTACTCAAGGATGACACCCAAATTAAGACAAAGTGAATGCTCCCC 1380
Qy 1388 TTCAGATCACACAGACAGCTACCGTGGCCTTGTATGATGATTTTGAATCATTTG 1447
Db 1381 TTCAGATCACACAGACAGCTACCGTGGCCTTGTATGATGATTTTGAATCATTTG 1440
Qy 1448 AGGCGCAAGAACTCTTATGATGAAATATCTTTGAAGTCCCTGACAGAAAGAAAGTGCCCA 1507
Db 1441 AGGCGCAAGAACTCTTATGATGAAATATCTTTGAAGTCCCTGACAGAAAGAAAGTGCCCA 1500
Qy 1508 TTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1567
Db 1501 TTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Qy 1568 AGCCTCAAAAGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1627
Db 1561 AGCCTCAAAAGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Qy 1628 GCTTTGAGAGGAGTGGAGTGTACCTCAGATGCTGCTCCTTCCATATATCTATTT 1687
Db 1621 GCTTTGAGAGGAGTGGAGTGTACCTCAGATGCTGCTCCTTCCATATATCTATTT 1680
Qy 1688 ATTACAACTCTTATATGAAAACATCAGTCTTACCTGAGAGATGATGATGATGATG 1747
Db 1681 ATTACAACTCTTATATGAAAACATCAGTCTTACCTGAGAGATGATGATGATGATG 1740
Qy 1748 AATTAGCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1807
Db 1741 AATTAGCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
Qy 1808 GGGAAAGAGACTCAAGCTTTGAGAGAGAAAGCAAAAGTATGCAATTTGCAAGAGCA 1867

```

```

Db 1801 GGGAAAGAGACTCAAGCTTTGAGAGAGAAAGCAAAAGTATGCAATTTGCAAGAGCA 1860
Qy 1868 TTTTGAAGAGACCCCGCATCTATCTATGATGAAGCTACTTCTGTTAGATTCGATTA 1927
Db 1861 TTTTGAAGAGACCCCGCATCTATCTATGATGAAGCTACTTCTGTTAGATTCGATTA 1920
Qy 1928 CTGAAGAGACTATTTCTTGTGTCATGAAGATGTTGTCAACACAGAACTTTTATTTTCA 1987
Db 1921 CTGAAGAGACTATTTCTTGTGTCATGAAGATGTTGTCAACACAGAACTTTTATTTTCA 1980
Qy 1988 TTGCAACACAGTTTCTTCAACAGTGTGATGATGATGAATTAATCTTGTGATGATGAGGTA 2047
Db 1981 TTGCAACACAGTTTCTTCAACAGTGTGATGATGATGAATTAATCTTGTGATGAGGTA 2040
Qy 2048 AGTAGCCGAACGTGTATACCAACCATGTTGCTTGTCAACCCCTCATATGATTCATTCAG 2107
Db 2041 AGTAGCCGAACGTGTATACCAACCATGTTGCTTGTCAACCCCTCATATGATTCATTCAG 2100
Qy 2108 AAATGTGCAATACACAGAGCAGCGGTGTCAGAACATGATMACCCCAATGGAGACAA 2167
Db 2101 AAATGTGCAATACACAGAGCAGCGGTGTCAGAACATGATMACCCCAATGGAGACAA 2160
Qy 2168 AGAAAGAAAATATATCCAAAGAGAGAGAAAGAAAGAACTTCAAGAAATTTGCAATA 2227
Db 2161 AGAAAGAAAATATATCCAAAGAGAGAGAAAGAAAGAACTTCAAGAAATTTGCAATA 2220
Qy 2228 GTGTGAAGAGGCTGTGGAACCTTGTGCTAAGTACATATGAGATTTTCTTTTGT 2286
Db 2221 GTGTGAAGAGGCTGTGGAACCTTGTGCTAAGTACATATGAGATTTTCTTTTGT 2278

```

```

RESULT 3
AAA62411
ID AAA62411 standard; cDNA; 2407 BP.
XX
AC AAA62411;
XX
DT 13-NOV-2000 (first entry)
XX
DE Human ATP-binding cassette transport protein nucleotide sequence.
KW Human; ATP-binding cassette transport protein; ABCxH; cytosolic;
KW natriuretic; neuroprotective; cerebroprotective; antidepressant;
KW anticonvulsant; antiparkinsonian; neuroleptic; cancer; epilepsy;
KW neurodegenerative disorder; akathisia; amorphous lateral sclerosis;
KW bipolar disorder; cataplexy; dementia; depression; Down's syndrome;
KW tardive dyskinesia; dystonia; multiple sclerosis; neurofibromatosis;
KW schizophrenia; Tourette's disease; ss.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT CDS /.*2252
FT /lag= a
FT /product= "ABCxH"
PN US6080842-A.
XX
PD 27-JUN-2000.
XX
PF 18-NOV-1998; 98US-0195391.
XX
PR 17-JUL-1997; 97US-0895522.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Shah P, Corley NC, Hillman JL;
XX
DR WPI: 2000-451228/39.
DR P-PSDB; AAB13355.
XX
PT New human ATP-binding cassette transport protein useful for diagnosing,
PT preventing and treating cancers, e.g. leukemia, sarcoma, and neuronal

```

PT disorders, e.g. schizophrenia, dementia -
XX
PS
XX Example 5; Fig 1; 31bp; English.

CC The present sequence encodes a human ATP-binding cassette transport
CC protein (ABCC8). The nucleotide sequence was isolated from a human
CC ovarian tissue cDNA library. Clones from the library were sequenced and
CC used as query sequences against homology databases. ABCC8 encoding
CC polynucleotides were extended using PCR. ABCC8 polynucleotides and
CC polypeptides may be used for the diagnosis, prevention and treatment of
CC conditions or disorders associated with the expression of ABCC8, e.g.
CC adrenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma,
CC teratocarcinoma, akathisia, Alzheimer's disease, amnesia, amyotrophic
CC lateral sclerosis, bipolar disorder, cataplexy, cerebral neoplasms,
CC dementia, depression, Down's syndrome, tardive dyskinesia, dystonias,
CC epilepsy, Huntington's disease, multiple sclerosis, neurofibromatosis,
CC Parkinson's disease, paranoid psychoses, schizophrenia, and Tourette's
CC disease.

XX
XX Sequence 2407 BP; 726 A; 477 C; 570 G; 629 T; 5 other;

Query Match 95.9%; Score 2248; DB 21; Length 2407;

Best Local Similarity 99.4%; Pred. No. 0; Mismatches 13; Indels 1; Gaps 1;

Matches 2265; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 8 TGTCTGGGATGATCTTTGGGCTGGGCGCGCGCGCTGCTTTCGAAAAGCGCGC 67
DB 1 TGTCTGGGATGATCTTTGGGCTGGGCGCGCGCGCTGCTTTCGAAAAGCGCGC 60
QY 68 ACTCGCGATTTCTGATCCGCGCTTATGCTCTGTTAGCGGCTCAGGTCCGATGAGGC 127
DB 61 ACTCGCGATTTCTGATCCGCGCTTATGCTCTGTTAGCGGCTCAGGTCCGATGAGGC 120
QY 128 CACATCAACTCGGCGCTTGGGAAACCGCTCGAGCTACCAATTCAGAGCTTAATAAA 187
DB 121 CACATCAACTCGGCGCTTGGGAAACCGCTCGAGCTACCAATTCAGAGCTTAATAAA 180
QY 188 GTATCACAATGCGAGATTTGGGAAAGGCAATTCAGAGCTTCTAGATCTGCAAGG 247
DB 181 GTATCACAATGCGAGATTTGGGAAAGGCAATTCAGAGCTTCTAGATCTGCAAGG 240
QY 248 CTCTCCAGGTATGCGCACTGATAGAAAAGAGACATGTTGCGATGCTCATCAGAGAG 307
DB 241 CTCTCCAGGTATGCGCACTGATAGAAAAGAGACATGTTGCGATGCTCATCAGAGAG 300
QY 308 GACTCCACAGACCCCAAAAGAGGTTAAAGATGTTGATCTCGGAAATCATTAAG 367
DB 301 GACTCCACAGACCCCAAAAGAGGTTAAAGATGTTGATCTCGGAAATCATTAAG 360
QY 368 CAATGCTTTCTTATGTGCGCCAAAAGACAGCCAGATCTACAGAGTAGAGTTCCATT 427
DB 361 CAATGCTTTCTTATGTGCGCCAAAAGACAGCCAGATCTACAGAGTAGAGTTCCATT 420
QY 428 CGCTGGAGTTTGGGCTGTCGAAAAGCCATGAAATATTTGCTTCCCTCATGTTTAAT 487
DB 421 CGCTGGAGTTTGGGCTGTCGAAAAGCCATGAAATATTTGCTTCCCTCATGTTTAAT 480
QY 488 ATGCTGAGACAGGCTCAACAGATGTCGGAACATGCTGAACCTAGAGATGACACAA 547
DB 481 ATGCTGAGACAGGCTCAACAGATGTCGGAACATGCTGAACCTAGAGATGACACAA 540
QY 548 ATACAGTTGCAACCATGCGACACAGAGTTCTGATTTGGCTATGATCAAGAGCTGAG 607
DB 541 ATACAGTTGCAACCATGCGACACAGAGTTCTGATTTGGCTATGATCAAGAGCTGAG 600
QY 608 CTGCTTTTATTAAGAAAGTTGAAATGCAATTTGGCAAGTAGCCCAAGATTCAATCC 667
DB 601 CTGCTTTTATTAAGAAAGTTGAAATGCAATTTGGCAAGTAGCCCAAGATTCAATCC 660
QY 668 GAAATATAGCCAAAAGTTCTTTCTCATCTTCAACACTGAGATCTGGGTTTCACTGA 727
DB 661 GAAATATAGCCAAAAGTTCTTTCTCATCTTCAACACTGAGATCTGGGTTTCACTGA 720

QY 728 GCAGACAGACGGAGCTTTATCTAAGGCTATTGAACAGAGAACAGGGGTATCAGTTTG 787
DB 721 GCAGACAGACGGAGCTTTATCTAAGGCTATTGAACAGAGAACAGGGGTATCAGTTTG 780
QY 788 TCCGATATGCTTTGGATTTTAATCTTCTCCATCACTTTGAAGTATGCTTGTCAAGT 847
DB 781 TCCGATATGCTTTGGATTTTAATCTTCTCCATCACTTTGAAGTATGCTTGTCAAGT 840
QY 848 GTGTTTGTATTAACAAATGCGGTCGCCAGTTTGCTTTGATACCTTGGAAACCTTGTA 907
DB 841 GTGTTTGTATTAACAAATGCGGTCGCCAGTTTGCTTTGATACCTTGGAAACCTTGTA 900
QY 908 CATACACAGATTCACAGTTTCAAGTCAACAGGTGAGAACTAATTTAGATAGAAATGA 967
DB 901 CATACACAGATTCACAGTTTCAAGTCAACAGGTGAGAACTAATTTAGATAGAAATGA 960
QY 968 ACAAGAGATTAATGATGACAGATTAATGCTGCTATAGCTCACTGCTGAATTAATGA 1027
DB 961 ACAAGAGATTAATGATGACAGATTAATGCTGCTATAGCTCACTGCTGAATTAATGA 1020
QY 1028 TGAAGTATTTAATTAATGAAGATTAATGAACACAGAGATATGATGATTTTGAAGAGT 1087
DB 1021 TGAAGTATTTAATTAATGAAGATTAATGAACACAGAGATATGATGATTTTGAAGAGT 1080
QY 1088 ATGAGACTGCTTATGAAAAAGTACCTCTACTGCTATGCTGAACCTTTGCTCAAGTG 1147
DB 1081 ATGAGACTGCTTATGAAAAAGTACCTCTACTGCTATGCTGAACCTTTGCTCAAGTG 1140
QY 1148 CTATTTTCACTGTCGTTTAAACAGCTTAATGCTGCTCCGACAGTCAAGGAATGTGAC 1207
DB 1141 CTATTTTCACTGTCGTTTAAACAGCTTAATGCTGCTCCGACAGTCAAGGAATGTGAC 1200
QY 1208 GTAACCTTACTGTTGAGATCTAGTAATGCTGAATGAGTCTTTTCACTTCAATTAC 1267
DB 1201 GTAACCTTACTGTTGAGATCTAGTAATGCTGAATGAGTCTTTTCACTTCAATTAC 1260
QY 1268 CCTGAACTTCTGCGAACTGTATATAGAGACTAGACAGACACTATGATATGAAACA 1327
DB 1261 CCTGAACTTCTGCGAACTGTATATAGAGACTAGACAGACACTATGATATGAAACA 1320
QY 1328 CTTTGTATCTTACTCACTCAAGTGAACACCAATTTAAAGCAAAAGTATGCTTCCC 1387
DB 1321 CTTTGTATCTTACTCACTCAAGTGAACACCAATTTAAAGCAAAAGTATGCTTCCC 1380
QY 1388 TTGAGATCAACACAGACAGCTAACGCTGCTTGAATATGAGCAATTTGATATACATG 1447
DB 1381 TTGAGATCAACACAGACAGCTAACGCTGCTTGAATATGAGCAATTTGATATACATG 1440
QY 1448 AGGAGCAGAAAGTCTTATGTAATATCTTTGAAGTCCCTGACAGGAAAGAAAGTGCCA 1507
DB 1441 AGGAGCAGAAAGTCTTATGTAATATCTTTGAAGTCCCTGACAGGAAAGAAAGTGCCA 1500
QY 1508 TTGTAGAGGTAGTGGTCAAGGAAAGCAATATGAGGCTATTAATTTGCTTATG 1567
DB 1501 TTGTAGAGGTAGTGGTCAAGGAAAGCAATATGAGGCTATTAATTTGCTTATG 1560
QY 1568 AGCTCAAAAAGGATGATTTATCTTGTGCTGCTCAAAATATACAAAGATGAGCTGAAA 1627
DB 1561 AGCTCAAAAAGGATGATTTATCTTGTGCTGCTCAAAATATACAAAGATGAGCTGAAA 1620
QY 1628 GCTTGGAGAGGAGTGGAGTGTACTCAGAGATCTGTCTTTCATTAATATAATTT 1687
DB 1621 GCTTGGAGAGGAGTGGAGTGTACTCAGAGATCTGTCTTTCATTAATATAATTT 1680
QY 1688 ATTAACAACCTCTTAATATGAAACATATATGCTTCACTGAGAAAGTATGACGTGCAA 1747
DB 1681 ATTAACAACCTCTTAATATGAAACATATATGCTTCACTGAGAAAGTATGACGTGCAA 1740
QY 1748 AATTAGTGACCTTCAATGATGCAATCTTCAATGCAATGCAATGCAATGCAATGCAATG 1807
DB 1741 AATTAGTGACCTTCAATGATGCAATCTTCAATGCAATGCAATGCAATGCAATGCAATG 1800
QY 1808 GGGAGAGAGACTCAAGCTTTCAGAGAGAGAAAGCAAGATGATGCAATTCAGAGGCCA 1867

```

|||||
Db 1801 GGAACGAGAGCTCAAGCTTTGAGGAGGAGAAAGCAAGAGTACCAATTCAGAGAGCA 1860
Oy 1868 TTTTGAAGAGACCCCGACGATACCTTATGATGAAGACTCTTCAATCGTTAGATTGATTA 1927
Db 1861 TTTTGAAGAGACCCCGACGATACCTTATGATGAAGACTCTTCAATCGTTAGATTGATTA 1920
Oy 1928 CTGAAGAGACTATCTTGGTGCATGAAGATGTGGTCAAAACAGAACTTCTATTTTCA 1987
Db 1921 CTGAAGAGACTATCTTGGTGCATGAAGATGTGGTCAAAACAGAACTTCTATTTTCA 1980
Oy 1988 TTGCACACAGATTGTCAACAGTGGTTGATGACAGATGAATTCATGTCTTGGATCAGGGA 2047
Db 1981 TTGCACACAGATTGTCAACAGTGGTTGATGACAGATGAATTCATGTCTTGGATCAGGGA 2040
Oy 2048 AGGTAGCCGAAAGCTGTGATCCCAACATGTTGCTTGAACCTCAATGATCTTATTCAG 2100
Db 2041 AGGTAGCCGAAAGCTGTGATCCCAACATGTTGCTTGAACCTCAATGATCTTATTCAG 2100
Oy 2108 AATGTGGCATACACAGAGCAGCCGTTGTGCAGAACCATGATACCCCAATGGAGCA 2167
Db 2101 AATGTGGCATACACAGAGCAGCCGTTGTGCAGAACCATGATACCCCAATGGAGCA 2160
Oy 2168 AGAAAGAAAATATATCCAAAGAGAGAAAAGAAACTACAGAGAAATTTGCAATA 2227
Db 2161 AGAAAGAAAATATATCCAAAGAGAGAAAAGAAACTACAGAGAAATTTGCAATA 2220
Oy 2228 GTGTGAAAAGCTGTGAAAACCTGTGCTAGTACATTAACAATTTCTTTTTCG 2286
Db 2221 GTGTGAAAAGCTGTGAAAACCTGTGCTAGTACATTAACAATTTCTTTTTCG 2278

```

RESULT 4

AAH14279 standard; cDNA; 2193 BP.

AAH14279;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:11609.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUN-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

full-length cDNAs -

Claim 8; SEQ ID 11609; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 2193 BP; 621 A; 451 C; 536 G; 585 T; 0 other;

Query Match 91.7%; Score 2150.4; DB 22; Length 2193;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2164; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Oy 1 ATGGCGTCTGCTCGGATGCAATTTCTTGGCGGCGCGCGGCGGCTTTGGAAAAG 60

Db 26 ATGGCGTCTGCTCGGATGCAATTTCTTGGCGGCGCGCGGCGGCTTTGGAAAAG 85

Oy 61 CGCGGCACTCCGAGATCTGATCCGGCTTTAGCTCTGTTAGCGGCTCAGTCCGAG 120

Db 86 CGCGGCACTCCGAGATCTGATCCGGCTTTAGCTCTGTTAGCGGCTCAGTCCGAG 145

Oy 121 TGAAGGCCATCAACTCGGCGCCTTGGAACCGCTCGAGCTTAC--CAGATTCAGAG 177

Db 146 TGAAGGCCATCAACTCGGCGCCTTGGAACCGCTCGAGCTTAC--CAGATTCAGAG 205

Oy 178 TCATTAATAAGTATCACTGCGAGAGATTGGGAAAAGGCAATTCAGACAGTTCTTAGAT 237

Db 206 TCATTAATAAGTATCACTGCGAGAGATTGGGAAAAGGCAATTCAGACAGTTCTTAGAT 265

Oy 238 GGTGCAAGGCTCTCCAGGTATGCGCACTGATAGAAAAGAGACATGTTGCAATGTCAT 297

Db 266 GGTGCAAGGCTCTCCAGGTATGCGCACTGATAGAAAAGAGACATGTTGCAATGTCAT 325

Oy 298 GCAGAGAGAGACTCCACACAGACCCAAAGAGGTTAAAGATGTTGATCTCGAAA 357

Db 326 GCAGAGAGAGACTCCACACAGACCCAAAGAGGTTAAAGATGTTGATCTCGAAA 385

Oy 358 ATCATTAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGCCAGATCTACAGCTAGA 417

Db 386 ATCATTAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGCCAGATCTACAGCTAGA 445

Oy 418 GTTGCAATTCGCTGGAGATTTTGGGTGTCGAAGGCGCATGAATATGTTGCTCCCTTC 477

Db 446 GTTGCAATTCGCTGGAGATTTTGGGTGTCGAAGGCGCATGAATATGTTGCTCCCTTC 505

Oy 478 ATGTTTAATATGCTGTAGACAGCTCAACAGATGTGCGGAAAATGCTGAACTGAGT 537

Db 506 ATGTTTAATATGCTGTAGACAGCTCAACAGATGTGCGGAAAATGCTGAACTGAGT 565

Oy 538 GATGACCAAAATACAGTTGCAACCATGCAACAGAGTTCTGATTTGCTATGCTATCA 597

Db 566 GATGACCAAAATACAGTTGCAACCATGCAACAGAGTTCTGATTTGCTATGCTATCA 625

Oy 598 AGAGCTGAGAGCTTTTAAAGAGAGTGAAGATGAGATTTGGCAAGTGGCCAG 657

Db 626 AGAGCTGAGAGCTTTTAAAGAGAGTGAAGATGAGATTTGGCAAGTGGCCAG 685

Oy 658 AATCAATCCAGAGATAGCCAAAATGTCCTTCATCTTCACTTCAACCTGATCTGGGT 717

Db 686 AATTCATCCAGAGATAGCCAAAATATGCTTCTCCATCTTCAACATCTGATCTGGGT 745
 Qy 718 TTTCACCTGACAGACAGACGGAGCTTTATCTAAGCTATGACAGAGAAAGGGGT 777
 Db 746 TTTCACCTGACAGACAGAGGGAGCTTTATCTAAGCTATGACAGAGAAAGGGGT 805
 Qy 778 ATCAGTTTGTCTGAGTGTCTTTGGTATTAATCTTCTTCCATCAGTTTGAAGTATG 837
 Db 806 ATCAGTTTGTCTGAGTGTCTTTGGTATTAATCTTCTTCCATCAGTTTGAAGTATG 865
 Qy 838 CTGTGAGTGTGTGTTTGTATTAACAATGCGGTGCCGTTTGGTCTTGGTAAACCTTGA 897
 Db 866 CTGTGAGTGTGTGTTTGTATTAACAATGCGGTGCCGTTTGGTCTTGGTAAACCTTGA 925
 Qy 898 ACACCTGTATACATACACAGCATTCACAGTTGACATCAGCGGTGAGAACTAGATTAGA 957
 Db 926 ACACCTGTATACATACACAGCATTCACAGTTGACATCAGCGGTGAGAACTAGATTAGA 985
 Qy 958 ATGAAATGAACAAAGCAGATTAATGATGACAGTAACTGCTATAGACTCATCTGTAAT 1017
 Db 986 ATGAAATGAACAAAGCAGATTAATGATGACAGTAACTGCTATAGACTCATCTGTAAT 1045
 Qy 1018 TATGAACTGTGAAGTATTTTATATATGAAGATATGAGACAGAGATATGATGATTT 1077
 Db 1046 TATGAACTGTGAAGTATTTTATATATGAAGATATGAGACAGAGATATGATGATTT 1105
 Qy 1078 TTGAAGAGTATGAGAGCTGCTTCACTGTAAGAACTGCTTCACTGCGGTATGCTGTAACCTT 1137
 Db 1106 TTGAAGAGTATGAGAGCTGCTTCACTGTAAGAACTGCTTCACTGCGGTATGCTGTAACCTT 1165
 Qy 1138 GGTCAAAAGTGTATTTTCACTGCTGCTTTTAAACAGTATATGCTGCTGCGCATGCAAGGGA 1197
 Db 1166 GGTCAAAAGTGTATTTTCACTGCTGCTTTTAAACAGTATATGCTGCTGCGCATGCAAGGGA 1225
 Qy 1198 ATTGTGCGAGGTATACCTTTACTGTTGGAGATAGTAATGCTGAATGCACTGCTTTTCAAG 1257
 Db 1226 ATTGTGCGAGGTATACCTTTACTGTTGGAGATAGTAATGCTGAATGCACTGCTTTTCAAG 1285
 Qy 1258 CTTTCATTAACCCCTGAACTTTCTGGAACCTGTATATGAGAGATAGTACAGCAAGCTGATA 1317
 Db 1286 CTTTCATTAACCCCTGAACTTTCTGGAACCTGTATATGAGAGATAGTACAGCAAGCTGATA 1345
 Qy 1318 GATATGAACACCTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1377
 Db 1346 GATATGAACACCTTGTACTCTACTCAAGGTAGACACCCAAATTAAGAACAAGTATG 1405
 Qy 1378 GCATCTCCCTTCAAGTACACCAAGACAGCTACCGTGGCTTTGATATATGTCATTTT 1437
 Db 1406 GCATCTCCCTTCAAGTACACCAAGACAGCTACCGTGGCTTTGATATATGTCATTTT 1465
 Qy 1438 GAATACATTTGAGGGGCAAGAAAGTCTTATGAGATATCTTTGAAGTCCCTGCGAGGAAG 1497
 Db 1466 GAATACATTTGAGGGGCAAGAAAGTCTTATGAGATATCTTTGAAGTCCCTGCGAGGAAG 1525
 Qy 1498 AAAGTGCATTTAGAGAGTATGAGTCAAGGAAAGCACAATATGAGAGCTTATTTT 1557
 Db 1526 AAAGTGCATTTAGAGAGTATGAGTCAAGGAAAGCACAATATGAGAGCTTATTTT 1585
 Qy 1558 CGCTTCTATGAGCTCAAAAAGGTAGCATTTATCTTGTGCTCAAAATATATCAAGATGTG 1617
 Db 1586 CGCTTCTATGAGCTCAAAAAGGTAGCATTTATCTTGTGCTCAAAATATATCAAGATGTG 1645
 Qy 1618 AGCCTGAAAGCCTTGGAGAGGCGAGTGGTATCTAGATAGTGTCTCTTCCAT 1677
 Db 1646 AGCCTGAAAGCCTTGGAGAGGCGAGTGGTATCTAGATAGTGTCTCTTCCAT 1705
 Qy 1678 AATATATTTATTAACAACCTTTATATATGAACAATCATGCTCTTCACTGAGAAAGTAT 1737
 Db 1706 AATATATTTATTAACAACCTTTATATATGAACAATCATGCTCTTCACTGAGAAAGTAT 1765
 Qy 1738 GCAGTGGCAAAATTAAGTGTGATTCATGATGCAATTTCTTGAATGCCAATGATATGAC 1797

Db 1766 GCAGTGGCAAAATTAAGTGTGATTCATGATGCAATTTCTGAAATGCCACATGATATGAC 1825
 Qy 1798 ACCCAAGTAGAGGAACGAGAGCTCAAGCTTTTCAAGAGAGAGAAAAAGCAAGTAGCAATT 1857
 Db 1826 ACCCAAGTAGAGGAACGAGAGCTCAAGCTTTTCAAGAGAGAGAAAAAGCAAGTAGCAATT 1885
 Qy 1858 GCAAGAGCCATTTTGAAGAGCCCCCAGTCATACCTATGATGAAGCTATCTTATGCTTGA 1917
 Db 1886 GCAAGAGCCATTTTGAAGAGCCCCCAGTCATACCTATGATGAAGCTATCTTATGCTTGA 1945
 Qy 1918 GATTGATTACTGAAGAGATATTTCTTGTGTCATGAAGAGATGCTGTCAAACAGAACT 1977
 Db 1946 GATTGATTACTGAAGAGATATTTCTTGTGTCATGAAGAGATGCTGTCAAACAGAACT 2005
 Qy 1978 TCTATTTTCACTGACACAGATATGTCACAGTGTGATGATGATGATGATGATGATGATGATG 2037
 Db 2006 TCTATTTTCACTGACACAGATATGTCACAGTGTGATGATGATGATGATGATGATGATGATG 2065
 Qy 2038 GATCAGGGTAAAGTAAAGCGGAGCGTGTACCCACATGCTTGTGCTTAACCTCATAGT 2097
 Db 2066 GATCAGGGTAAAGTAAAGCGGAGCGTGTACCCACATGCTTGTGCTTAACCTCATAGT 2125
 Qy 2098 ATCTATTTCAAGAAATGTGGCATACACAGAGAGCGGTGTGAGAAACCATGTAACCCCAA 2157
 Db 2126 ATCTATTTCAAGAAATGTGGCATACACAGAGAGCGGTGTGAGAAACCATGTAACCCCAA 2185
 Qy 2158 TGGGAAGC 2165
 Db 2186 TGGGAAGC 2193

RESULT 5
 ID ABL01991
 ABLO1991 standard; cDNA; 2182 BP.
 XX AC ABL01991;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 455.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PA 11-JUL-2000; 2000US-0614150.
 XX PE (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PMD, Myers BW;
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB57888.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 455; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pcf_sequences](http://wipo.int/pub/published_pcf_sequences).

XX
 SQ Sequence 2182 BP; 491 A; 588 C; 633 G; 470 T; 0 other;

Query Match 24.1%; Score 564.8; DB 23; Length 2182;
 Best Local Similarity 58.5%; Pred. No. 5,9e-152;
 Matches 1011; Conservative 0; Mismatches 702; Indels 15; Gaps 1;

```

Qy 354 GAAATTCATTAAGCAATCTTTATGTGTCGCCAAGACAGAGCCAGATCTACGAGC 413
Db 358 GGAATGCGCGCCATCATATGCGCTACATCTGCGAAGAGAGATCCACTGCTGGAAA 417
Qy 414 TAGAGTTCATTTGCGTGGGATTTTGGGTGTGCAAGGCCATGAATATTGTGGTCC 473
Db 418 GCGGGTGGGATTTTCCCTGGGTCTGTGGCTGCTCAAGCTGCTGACCGCTGTGCTCC 477
Qy 474 CTTCAATTTAATATGCTGTAGACGCTCAACGAGATGTGGGAAACATGCTGAACCT 533
Db 478 CTTCTGTTCAAGAGAGCTGTGACAC-----CGATGACACGCTGAACAT 522
Qy 534 GAGTATGACCAAAATACAGTTGCAACCATGCAACAGAGATTTGATGTGCTATGCTGT 593
Db 523 GGAACATGCGCCGAGTGGGCTCTCCGCAACACCGACATGATGTGGATATGAGCAT 582
Qy 594 ATCAAGAGCTGAGCTGCTTTTATTAAGAGTTGCAAAATGCAATATTGGCAAGTAC 653
Db 583 TGCTGAGCGCATGGCGCGGCTTTAATGATGGCGAATGCAAGTGTGGCCAAAGTGGC 642
Qy 654 CCAGAAATCAATCCGAAGATAGCCAAAGTCTTTCCATCTTCACTCAACCTGGATCT 713
Db 643 CCACCACTGATCCGAAGATGCGACAGCATGTCTGCAATCTGCAACACTGATCT 702
Qy 714 GGGTTTCACTGAGCAGACAGACGGAGCTTTATCTAAGCTATTGACAGAGAAACAG 773
Db 703 GGCCTTCCACCTGGAACAAACAACTGAGCACTGTCAAAAGACATGATCGAGATCGAG 762
Qy 774 GGGTATACGTTTGTCTGAGTCTTGGATTTAACTCTTCCCATCATGTTGAAGT 833
Db 763 GGGCATTTAACTTTGTGCTCTCCGCAATGCTTTCAACATTTGCCACATCTTTGAGCT 822
Qy 834 GATGCTGTCAGTGTGCTTTGATTAACAATGCGGTGCGCCAGTTGCTTGTGAACCT 893
Db 823 GGCCTCTGTCAGATATCTGGAGTGAAGTGTGGCTGCGCTTGTGCTGACGAT 882
Qy 894 TGGAACTTTGTTACATACAGACATTTCAAGTTGCACTCAACGGTGAGAACTAGATT 953
Db 883 GGGCGCGTGGCATATAGCTGCTCACTGAGTGTGACCCAGTGGCGACCCAGTT 942
Qy 954 TAGAATAGAATGAACAAGACAGATATATGATGAGGATATGCTGTATAGACTCACTGCT 1013
Db 943 CGGTGATTTATGAACCAAGCGGAGAACGAGCGCGCAACAAGCGCTGAGATCTCGCTGAT 1002
Qy 1014 GAATTAATGAATCTGTAATATTTTAATGAAGATATGAAGACAGAGATATGATG 1073
Db 1003 CAACTAACGAGAGGTGAATTAATCTTCAACAGAAAGTACGAGCGGGTGTGCTACACGA 1062
Qy 1074 ATTTTGAAGAGATGAGATGCTTCAATTTGAAAAGTACCTTACTGCTGATGCTGAA 1133
Db 1063 GGTGCTGAAGAGATAGAGCGCGCAGCTTGAAGACTACTCAAGTCTGCTGTCTCA 1122
Qy 1134 CTTTGGTCAAGTCTATTATTAAGTGTGCTTTACAGCTATATATGCTGCTGCAAGTCA 1193
Db 1123 CTTTGGCCGAAAGCGCATTTTCAAGCGCACAGTGTGATTAATGATGATTTGGCCGCA 1182
Qy 1194 GGGAAATGTGAGAGTACCTTACTGTTGAGATCTAGAAATGGAATGAGACTGCTTTT 1253
Db 1183 GGAAGATTGCCAAGGCAACATGACGGTGTGAATTTGGTATGATGCTCAAGCCCTGCTTT 1242

```

```

Qy 1254 TCAGCTTTCAATACCCTTGAACTTTCTGGAACTGTATATAGAGACTAGACAGACT 1313
Db 1243 CCAGCTCCGATCCCCCTCGGCTTTCTGGGCAAGTTTATCTGATGAGTGGCAGCGGCTTT 1302
Qy 1314 CATGATATGAACCTTGTTTACTTACTCAAGTATGACCCCAATTAAAGCAAGT 1373
Db 1303 GCTGACATGCGGCGCATGTTCACGTTGATGAAGTGAACGTGACGTGACATTCAGACGCC 1382
Qy 1374 GATGCACTCTCCCTTCAGATCACACACAGACACTACCGTGGCTTTGATTAATGCA 1433
Db 1363 CAATGCCAGGCCCTGTTCTGTGACACACACTAATCATCATTTAGTTCCGCAAGTGAG 1422
Qy 1434 TTTTGAATACATTAGAGCGCAAGAAATCTTTATGTAATATCTTTGAATGCTCTGCAAG 1493
Db 1423 CTTGAGATACGAACCTTGCAAGCCCATTTTCCGGAGCTCAGTTTCAACATACCCGCGG 1482
Qy 1494 AAAAGAAAGTGGCCATTTGATGAGATGAGTGTGCGGAAAAAGCAGATGAGGCTATT 1553
Db 1483 TAAAAAGTCCCATTTGTGGGCGGCTCTGCTCAGAAAAATCGTATGATGCGCTTTCT 1542
Qy 1554 ATTTGCTTTATGAGCTCAAAAGGATGATTTATCTTGTGTCMAAATATACAGA 1613
Db 1543 CTTGCGCTTTTATGAGCCAAATCTGTGTAAGTGTGATGCTGCGCAGATATCAGCGC 1602
Qy 1614 TGTAGCCTTGAAAGCCTTGGAGGCAATGAGAGTGTATCTCAGATGCTGTCTTT 1673
Db 1603 CGTGGATTTGAGAGCCTGCGCAAGTTATGTCAGTGTGTCGCAAGATTCGCTGATTT 1662
Qy 1674 CCATATATCTATTTATTAACAATCTTATATGAAACATCAGTCTTCACTGAGGAAGT 1733
Db 1663 CCACAAACCATGACAGCAATATATCACTATGGAACCTGTCAAGTCCAGCGAAGT 1722
Qy 1734 GTATGAGTGGCAAAATTTAGCTGACTTATGATGCAATTTCTTGAAATGCCACATGATA 1793
Db 1723 GCAAGATGCTCCCGCATAGCTGTGATTTGCAATGATCATATGAGTGGCAGAGAGTA 1782
Qy 1794 TGAACCCCAATGAGGAGAGAGAGATCTAGCTTTCAAGAGAGAAAAAGCAAGATAGC 1853
Db 1783 CTCACATGAGGTAGCGCAACGTGATTAAGTTATCCGCTGCGCAGAGAGCGCCGTTC 1842
Qy 1854 AATTGCAAGACCAATTTTGAAGAACCCCGCATATCTGATGATGAAGCTACTTATC 1913
Db 1843 CATGCTAGGGCCATPACTGAAGAAACCCCATCTGATCTTTGACAGGCTTACAGCAG 1902
Qy 1914 GTTATGATCGATTAAGTGAAGACTATTTCTTGTGTCAGTGAAGAGTGTGCAACAAG 1973
Db 1903 TCTGATCTGATCAACAGCATTAATTTCTGACAGGCTGTGACCCGCGTACTCAGAG 1962
Qy 1974 AACTTCTAATTTTCAATGACACAGATTTGCAACAGTGTGATGATGCAATGATCTGT 2033
Db 1963 CACTAGCATTTTGCAATTTGCCACCGTCTCTGACAGTCAAGGATGCGGACGAGATTTGCT 2022
Qy 2034 CTTGATACGGGTAAAGTGAAGCCGAACGTGTGATCCCAATGATGTTGCT 2081
Db 2023 GCTGAGAACGAGACGTGATGATGAACGTGACCCCATTTGGAGGCTCT 2070

```

RESULT 6
 AB101989
 ID AB101989 standard; cDNA; 2539 BP.

XX AB101989;
 XX 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 449.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.

XX 27-SEP-2001.
PD
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW,
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB57886.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 449; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 2539 BP; 580 A; 677 C; 708 G; 574 T; 0 other;
SQ

Query Match 24.1%; Score 564.8; DB 23; Length 2539;
Best Local Similarity 58.5%; Pred. No. 6,4e-152;
Matches 1011; Conservative 0; Mismatches 702; Indels 15; Gaps 1;

QY 354 GAATATCATTAAGCAATGCTTCTTATGTGTGGCCCAAGAGCCAGATCTACGAGC 413
DB 715 GGAATAGCTGGTGCATGATGCTTACATCTGGCCAAAGAGATCCATGTCGAAA 774
QY 414 TAGAGTGGCCATTTGCTGGGATTTTGGGTGTGCAAGGCCATGATATGTGTCC 473
DB 775 GCGGGTGGGATTTTCCGTGGGTCTGTGGTGTGCTCCAGCTGTGACCGCTGTGCC 834
QY 474 CTTGATGTTTAATATGCTGTGAGACAGCTCAACCAAGTGTGGGAAACATGTAACCT 533
DB 835 CTTCTGTTCAAGAGAGCTGTGACA-----CGATGACACGCTGAACAT 879
QY 534 GAGTATGACACCAATACAGTTGCAACATGCAACAGCAAGTTCTGAGTGTGTGT 593
DB 880 GGAACACTGCCCCGAGTGGGCTCTCTCCGACGACCGCACTGATGTGGATATGTCAT 939
QY 594 ATCAAGAGCTGAGAGCTGCTTTTAAAGAAATGCAATGCAATGTTGGCAAGTAC 653
DB 940 TGCATAGAGCCAGTGGGCGGCTTAAATGAGTTGCGAATGCAATGTTGGCAAGTAC 999
QY 654 CCAGAAATTCATCCGAAGATAGCCAAAATGCTTTTCTCATCTTCAACCTGATCT 713
DB 1000 CCACACCTCGATCCGAAAGATCGCACCAATGTTCTCTGATCTGACAACTGATCT 1059
QY 714 GGGTTTCACTGAGCAGACGAGCGGAGCTTATCTAAGGCTTTGACAGAGGAAAG 773
DB 1060 GGGCTTCACTGAGCAGCAGCAAACTGAGCAGCTGCAAAAGCAATGATGAGATGAG 1119
QY 774 GGGATACAGTTTGTCTGAGTGTGATTTAATCTTCTTCCATCATGTTGAAGT 833
DB 1120 GGGCATTAACCTTGTGTCTCCGCAATGCTTCAACATGTTGCTCATCTTTGAGCT 1179
QY 834 GATCTTGTCAAGTGTGTTTGTATTAATAATGCGGTGCGCCAGTTGCTTGTGTAACCT 893

DB 1180 GGCCCTGTGTCCAGTATCTCTGGAGTGAAGTGTGGCTTGCCCTGCTGTGACAT 1239
QY 894 TGAACACTTGTGATACATACAGCATTCACAGTTGACATCAAGTGTGAGAACTAGATT 953
DB 1240 GGGCTGTGGTATACATACCTGTCTTACATCTGAGTGTGACCAAGTGTGGCAGCCAGTT 1299
QY 954 TAGAATGAAATGAAACAAAGCAGATTAATGATGCAATGCTATGATGATGCTTCT 1013
DB 1300 CCGTGTATTTATGAAACGAGCGGAGAAACGAGCGGCAACAGGCGGTAGACTGTGAT 1359
QY 1014 GAATTAAGAACTGTGAAGTATTTTAATATGAAGTATGAAGCAGAGATGATG 1073
DB 1360 CAATACGAGACGGTGAAGTATTCACAAAGAAAGTACAGCGGGTGTCTAACG 1419
QY 1074 ATTTTGAAGACGTATAGACTGCTTCAATGAAGAACCTTACTCTGGCTATGCTGAA 1133
DB 1420 GGTCTGAAGAAATATAGAGCGGCGAGCTTGAAGACTGATCCAGTTCGCTGTGCTCA 1479
QY 1134 CTTTGTCAAAAGTCTATATTTCAAGTGTGCTTTAACAAGCTATTAATGCTGCACTCA 1193
DB 1480 CTTTGGCCAGAACGCGCATTTTACAGACGCACTGAGCTGATTAATGATTTGGCCGCA 1539
QY 1194 GGAATGTGCGAGATACCTTACTGTGAGATCTAGTAAATGATGATGATGCTTTT 1253
DB 1540 GGAATGTGCGAGATACCTTACTGTGAGATCTAGTAAATGATGATGATGATGCTTTT 1599
QY 1254 TCAGCTTTCATTAACCTCTGAACCTTCTGGGACCTGTATATAGAGACTAGACAGCACT 1313
DB 1600 CAGCTCTCATCTCCCTCGCTGCTTGTGAGATGTTTATGAGGTGAGACAGGCTTT 1659
QY 1314 CATGATATGAAACACTTGTGTTACTTACTCAAGATGAGACCCAAATTAAGCAAAAGT 1373
DB 1660 GGTGACATCGGGGCCATGTTTCAAGTATGAGACAGTATGAGTATGACAGCGCG 1719
QY 1374 GATGCAATCTCCCTTCAATTCACACAGACAGCTACCGTGTGCTTATATGCTCA 1433
DB 1720 CAATGCGCAAGCCCTGTGTGTGACACACTACTCATTCATGATGATTCGCAACGTGAG 1779
QY 1434 TTTTGAATACATTTAGAGGCGAGAAAGCTTATGGAATATCCTTGAATGCTCCTGAGG 1493
DB 1780 CTTGAGATGAACCTGTGCAAGCCATTTTCCGGACCTAGTTTCAACATACCCGCGG 1839
QY 1494 AAGAAAGTGGCAATTTGTAGAGTATGAGTGTGAGGAAAGCAATATGAGGCTATT 1553
DB 1840 TAAAGAGTGGCAATTTGTAGAGTATGAGTGTGAGGAAAGTATGAGTGTGAGGCTTCT 1899
QY 1554 ATTTGCTTATAGAGCTTCAAAAGGTAGCAATTTATCTGTGCTCAAAATATACAGA 1613
DB 1900 CTTCCGCTTCTTGAAGCAAACTGTGTTAAAGTGTGATGCTGTGCGCAGATATACAGCGC 1959
QY 1614 TGTAGCTGTGAAAGCTTGTGAGGCGAGTGTGAGTGTGATCTCAGAGTGTGCTCTT 1673
DB 1960 CGTGATTTGGAAGCTGTGCAAGGTATGTGCAAGTGTGCGCAGAGTATGAGTGTGCTT 2019
QY 1674 CCATTAATCTATTTATTAACAACCTCTTATATGAAACATCAAGTGTCTTCACTGAGAA 1733
DB 2020 CCACAAACCATGAGACAAATATCACTATGSCAACTGTGCCAAGTCCACGCAAGAGT 2079
QY 1734 GTATGCAATGCGCAAAATATGCTGCACTTCAATGATGATATTTCTTGAATGCAATGATA 1793
DB 2080 GCAGAAATGCTGCGCGGATGCTGATTTGATGATGATGATGATGATGATGATGATGAT 2139
QY 1794 TGACACCAAGTATGAGGAGGAGAGAGTCAAGCTTTCAGGAGGAGGAGGAGGAGGAGTAC 1853
DB 2140 CTCACCTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 2199
QY 1854 AATTGCAAGGCAATTTTGAAGAGCCCGCAGTCACTATCTATATGAGGAGTACTTCTATC 1913
DB 2200 CATGCGTATGAGGCACTATGAGAGCAACCCCATCTTGATCTTTTGAAGAGGCTTCAAGAG 2259
QY 1914 GTTAAATGATTAATGAGAGAGCTATTTCTGCTGCTCAATGAGATGCTTCAACAGAG 1973
DB 2260 TCTGACATGATCACGAGATTAACATCTGTGAGGCTGTGAGCCGCGCTTACAGGAGG 2319

Qy 827 TTGAAGTATCTGTGACGTGCTGTTTGTATTAACAAATGCGGTGCCAGTTGCTTTGG 886
 Db 4109 TTGAGCTGGCCCTGCTGTCCAGTATCCCTGGAGTGAAGTGGCTGCTCCCTGCTG 4168
 Qy 887 TAACCTTGGAAACCTTGTCATACACAGATTCAACGTTGACGTACACAGGTGGAGAA 946
 Db 4169 TCACATATGGGCTGGCTGGCATATACCTGCTTACCTGAGTGTGACCCAGGGGCA 4228
 Qy 947 CTAGATTAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1006
 Db 4229 CCGAGTTCGGTATTTATGAAACGAGGCGAGAAAGAGGCGGCAACAGGCGGTAGCT 4288
 Qy 1007 CACTGCTGATTAAGAACTGTGAATTTTAAATGAAAGATGAGACACAGAT 1066
 Db 4289 CGCTGATCAACTACGAGCGGTGAAGTACTTCAACAAAGAAATACGAGCGGGTGGCT 4348
 Qy 1067 ATGATGATTTTGAAGCGTATGAGACTGCTTATGAAAGTAAGTACTTCTGCTA 1126
 Db 4349 ACAACGAGGTGCTAAGAAATGAGGCGGCGACCTGGAAGTACGCTCCGCTC 4408
 Qy 1127 TGCTGAACCTTGTCAAAAGTCTATTTTCACTGCTGCTTAAAGTAAATGCTGCTG 1186
 Db 4409 TGCTCAACTTTGGCGAGAAAGCCATTTTCAAGACGACCTGATGATTTGATTTGG 4468
 Qy 1187 CCACTCAGGGAATTTGGCAGGTAACCTTACTGTTGAGATCTAGTAATGCTGAATG 1246
 Db 4469 CCGCAAGGATTTGCCAAGGCAACATGACGTTGAGATTTGGTATGCTCAACGCC 4528
 Qy 1247 TGCTTTTACGCTTCAATTCCTGAACTTTCTGGGAACCTGTAATAGAGACTAGAC 1306
 Db 4529 TGCTTTTACGCTTCAATTCCTGAACTTTCTGGGAACCTGTAATAGAGACTAGAC 4588
 Qy 1307 AAGCACTATGATATGAAACCTTGTCTTACTCTCAAGGTAGACACCAATTAAG 1366
 Db 4589 AGGCTTTGCTGATGATGCGGCGCATGTTCACTGATGATGATGATGATGATGATG 4648
 Qy 1367 ACAAGTATGATGATCTTCCCTTCAATCACACACAGACGTAACCGTGGCTTTGATA 1426
 Db 4649 CAGCGGCAATGCCAGCGCTGTTCTGAGACACCACTCACTCATGATGATGATGATG 4708
 Qy 1427 ATGTCATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
 Db 4709 ACGTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4768
 Qy 1487 CTGCAAGAAAGAAAGTGGCCATTTGAGAGGTAGTGGCTGAGGAAAGCAATATGTA 1546
 Db 4769 CCGCGGTAAGAAAGTGGCCATTTGAGAGGTAGTGGCTGAGGAAAGCAATATGTA 4828
 Qy 1547 GGCCTATTTTGGCTTATGAGCTTCAAAAGGTAGCATTTATCTGCTGCTCAAAATA 1606
 Db 4829 GCTTCTCTTCCGCTTCTTTGAGCAAACTCTGTTAAAGTGTGATGCTGGTGGCAAGATA 4888
 Qy 1607 TACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1666
 Db 4889 TCAGCGCGGTGATTTGAGAGCTGCGCAAGGTTATTTGAGTGTGCTCCGAGGATTCGG 4948
 Qy 1667 TCCCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1726
 Db 4949 TGCATTTCCACACACATGAGACATATTCATATGCACTGCACTGCAAGTCCACAG 5008
 Qy 1727 AGGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1786
 Db 5009 CAGAAAGTACAGAAAGTCTGCCGATGCTGATTTGATGATGATGATGATGATGATG 5068
 Qy 1787 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1846
 Db 5069 GACAGTATCTCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 5128
 Qy 1847 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1906
 Db 5129 GCGTGGCATGCTGATGAGGCTATGATGAGAAACACCCCATTTCTGATCTTTGACAGGCTA 5188
 Qy 1907 CTTCATGTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1945

Db 5189 CAAGACTCTGAGCTGATCAGCAGAGCATGATGATG 5227
 RESULT 11
 AAA45541
 ID AAA45541 standard; cDNA, 437 BP.
 XX
 AC AAA45541;
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human secreted expressed sequence tag SEQ ID NO:2116.
 XX
 KW Human; mouse; chicken; rat; secreted expressed sequence tag; SESTR;
 KW expressed sequence tag; EST; probe; chemokine; analgesic; proliferative;
 KW immunomodulatory; haematopoietic; chemokine; analgesic; haematostatic;
 KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiaesthetic; vulnary; antiparkinsonian;
 KW anticancer; osteoprotective; neuroprotective; nociceptive; antiparkinsonian;
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
 KW tumour; infection; depression; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20021991-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 15-OCT-1999; 99WO-US24206.
 XX
 PR 15-OCT-1998; 98US-010436.
 XX
 PA (GENE) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR;
 XX
 DR WPI; 2000-317938/27.
 PT Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (SESTRs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -
 XX
 PS Claim 1; Page 731; 803pp; English.
 XX
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
 CC sequence tags (SESTRs), isolated from human, mouse, chicken and rat
 CC tissue sources. The SESTRs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemokine; analgesic; immunomodulatory; haematopoietic;
 CC chemokine; analgesic; immunomodulatory; haematopoietic;
 CC cytostatic; antibacterial; antileukemic; antidiabetic;
 CC antiaesthetic; vulnary; anticancer; osteoprotective; neuroprotective;
 CC nociceptive; antiparkinsonian; antipsychotic; cerebroprotective;
 CC anticonvulsant; and antidepressant. The SESTRs can be used for gene
 CC therapy and in vaccines. The SESTRs are useful as probes for the
 CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the SESTRs. Proteins encoded by the SESTRs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and

CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.

XX Sequence 437 BP; 119 A; 96 C; 101 G; 121 T; 0 other;

Query Match 17.4%; Score 408.6; DB 21; Length 437;
Best Local Similarity 99.0%; Pred. No. 2.7e-107;
Matches 411; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1096 GCTTATGAAAGTACCTCTACTGCTATGCTGAAGTCTTTC 1155

Db 21 GCTTATGAAAGTACCTCTACTGCTATGCTGAAGTCTTTC 80

Qy 1156 AGTGTGGTTTAAAGTATTAATGCTGCGAGTCAAGGAAATGGCAGTACCTT 1215

Db 81 AGTGTGGTTTAAAGTATTAATGCTGCGAGTCAAGGAAATGGCAGTACCTT 140

Qy 1216 ACTGTTGAGATCTAGTAATGCTGAATGAGTCTTTTTCAGCTTTCATTACCCCTGAAC 1275

Db 141 ACTGTTGAGATCTAGTAATGCTGAATGAGTCTTTTTCAGCTTTCATTACCCCTGAAC 200

Qy 1276 TTTCTGGGAACCTGTATATGAGAGTACAGACATCATATGAAACCTTGT 1335

Db 201 TTTCTGGGAACCTGTATATGAGAGTACAGACATCATATGAAACCTTGT 260

Qy 1336 ACTCTACTCAAGTACAGACCCCAATTAAAGCAAAAGTATGAGCATCTCCCTCAGATC 1395

Db 261 ACTCTACTCAAGTACAGACCCCAATTAAAGCAAAAGTATGAGCATCTCCCTCAGATC 320

Qy 1396 ACACCAAGACAGCTACCGTGCCTTTGTAATGTCATTTTGAATACATTGAGGCCAG 1455

Db 321 ACACCAAGACAGCTACCGTGCCTTTGTAATGTCATTTTGAATACATTGAGGCCAG 380

Qy 1456 AAAGTCTTATGAGTAATCTTTGAAGTCCCTGAGGAAAGAAATGCGCCATTG 1510

Db 381 AAAGTCTTATGAGTAATCTTTGAAGTCCCTGAGGAAAGAAAGGCGCTCG 435

RESULT 12

AAx78080
ID AAx78080 standard; DNA; 5369 BP.

AAx78080;

19-AUG-1999 (first entry)

A. goseypil ADE4 DNA.

XX Purine biosynthesis; phosphoribosylpyrophosphate synthetase; KPR1; KPR2;
KW glutamine-phosphoribosylpyrophosphate amidotransferase; ADE4; GUA1;
KW IMP-dehydrogenase; GUA1; GMP-synthetase; GUA2; PRPP synthetase; mouth;
KW riboflavin; vitamin B2; treatment; inflammation; throat; mucosal layer;
KW skin disorder; ss.

XX Ashbya goseypil.

XX Key Location/Qualifiers

FT CDS 55.1482

FT /tag= a

FT /product= "ORF1"

FT 1767..3299

FT /tag= b

FT /product= "ADE4 gene encoding glutamine PRPP

FT CDS 3588..4703

FT /tag= c

FT /product= "ORF2"

XX EP927761-A2.

XX 07-JUL-1999.

XX 08-DEC-1998; 98EP-0123331.

XX 23-DEC-1997; 97DE-1057755.

XX (BADI) BASF AG.

XX Hoeftlen HW, Jimenez A, Pompejus M, Revuelta Doval JL;

XX Santos Garcia MA, Seuburger H;

XX WPI; 1999-373465/32.

XX P-PSDB; AA08961, AA08962, AA08963.

XX Ashbya goseypil purine biosynthesis enzymes and related genes for

XX production of riboflavin

XX Claim 12; Page 15-22; 48pp; German.

CC This invention describes novel enzymes involved in purine biosynthesis

CC in Ashbya goseypil. The enzymes described in the invention include

CC phosphoribosylpyrophosphate (PRPP) synthetases KPR1 and KPR2

CC glutamine-phosphoribosylpyrophosphate amidotransferase (ADE4)

CC IMP-dehydrogenase (GUA1) and a GMP-synthetase (GUA2). The nucleic acid

CC sequences encoding the enzymes involved in purine biosynthesis are used

CC to genetically modify microorganisms, for the production of riboflavin.

CC Riboflavin, also known as vitamin B2, is essential in humans and animals

CC as well as used to treat inflammation of the mouth and throat mucosal layer

CC ADE4 gene which encodes a glutamine PRPP amidotransferase and 2 other

CC open reading frames.

XX Sequence 5369 BP; 1351 A; 1333 C; 1299 G; 1386 T; 0 other;

Query Match 15.3%; Score 358.6; DB 20; Length 5369;

Best Local Similarity 54.1%; Pred. No. 3.1e-97;

Matches 777; Conservative 0; Mismatches 69; Indels 9; Gaps 2;

Qy 703 AACCTGATCTGGTTTACCTTGAAGACAGACAGCGAGCTTATCTAAGCTTATGAC 762

Db 1 AAGCTTGACCTTGGCTGACCTTGAAGTGCAGACAGCTGAGTAAACCGCAATGAT 60

Qy 763 AGAGCAAGAGGATATGATTTGCTGAGTGTGCTTGTATTAATCTTCCATC 822

Db 61 CGTGTGTAAAGTATCTTATGTCAGTGCATAGTGTGTTTTCATATATCCGATT 120

Qy 823 ATGTTGAAGTATGCTTGCAGTGTGTGTTTGTATTAACAATCGGTGCCAGTTGCT 882

Db 121 ACATTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

Qy 883 TTGTAACCTTGAAGACCTTGTATCATACACAGCATTCACAGTTGACGTACAGGTG 942

Db 181 GCTATTAACATTCGACATATCTTCTTATCTCAATCTTCTTCTTCTTCTTCTTCTT 240

Qy 943 AGAAGTATTAAGTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1002

Db 241 CGACACGCTTTAGCGGTGAGTGAAGCAAGGCTCAATTAAGGCGCTAGTGTGCAATTG 300

Qy 1003 GACTCAGTCTGAATATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1062

Db 301 GATTCCTTAATTAATTTTGAAGCTGTAAAGTATTTCAATTAAGGAGTACCTTGGGAC 360

Qy 1063 AGATATGATGATTTTGAAGACGTATGAGCTGCTTCAATTAAGGAGTACCTTCTG 1122

Db 361 AAGTATGACATCTTGTATGAAGTACCGGATTCAGATTAAGGATTCGCAATCGCTG 420

Qy 1123 GCTATGCTGAATTTTGTCAAGTGTCTTTCAGTGTGCTTTCAGTGTGCTTTCAGTGT 1182

Db 421 GCGTTTGAACACCGGACAGAACTTAATTTTCACTGACGACGACGACGACGACGAC 480

Qy 1183 CTCGCCAGTCAAGGATATGTCAGGATACCTTCTGTTGAGATCTGATGATGATGAT 1242

Db 481 ATGCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

Qy 1243 GAGTCTTTTTCAGCTTCAATCCCTGAACTTTCGGAAGTATATGAGAGACT 1302

Dh 541 CAACCTGATTCACGCTCCGTGCACTAACTTCCTGTAGCGCTACCGTATCTC 600
Qy 1303 AGACAAGCACTCATATGATATGACACCTTGTCTTACTCAAGTAGACACCAATT 1362
Dh 601 AAGAGCTCTGATATGATATGAAATCTTTATTTAACTGCAAAAAATCAGCTCAATT 660
Qy 1363 AAAGACAAAGTAGGAGCTCTCCCTTCAGATCAACACAGACGACGAGCCCTT 1422
Dh 661 AAGAACTCCCAAAATGCCAGAACCTACCAATACAAACCGTTGATAT---TTCGCTTT 717
Qy 1423 GATATATGCACTTTGAAATATATGAGGCGCAAGATCCCTTATGGAATTCCTTGAA 1482
Dh 718 GAAATATGATATGATATGATATGACCCGAGCGGATATGAAATGTTTCTTTACC 777
Qy 1483 GTCCCTCAGAGAAAGAAAGTAGGAGCTATGAGAGTAGTGGGTGAGGAAAGACAATA 1542
Dh 778 ATCCAGCTGGAATGAAAGACCTGCAATAGCCCATCGGCTCGGGAATCCACATT 837
Qy 1543 GTGAGCTATTTATTCGCTTATGAGCTCAAAAGGTAGCATTTATCTTGTGTCAA 1602
Dh 838 TTGAAGCTCGATTTAGATCTATGAGCCGAGCAAGTCTATCTAGTTGGCGGACA 897
Qy 1603 AATATACAAAGTAGGAGCTGGAAGCCTTGGAGGAGAGGAGTGTACTCAGAT 1662
Dh 898 GATATCCGCGATTTAGACTGTCTTTTACGGAAGCTATCGGTGCGCCCAAGT 957
Qy 1663 GCTGCTCTCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1722
Dh 958 ACTCTCTCTTCAATGACAACTCGGAGAAATGTTAAATTCGCAATATATGTTCTCT 1017
Qy 1723 CTTGAGGAAGTATGACAGTGGCAAAATTAAGTGAATTCATGATGCAATTCCTGATG 1782
Dh 1018 GACGATGATTTCTCAGGGCCATGAAAAAGCTCACTCAGAACTACTCAGAACCTA 1077
Qy 1783 CCAATGATATGACACCCCAAGTAGGGAAGAGACGCTCAAGCTTTCAGAGGAGAAAG 1842
Dh 1078 CCAAGAGCGCTTCACCGTTGTAGGGAGCGCGTTGATGATCAAGGAGGTGAGAAA 1137
Qy 1843 CAAAGATGAGCAATTCGAAAGACCAATTTGAAGAGCCCGGAGTCACTACTATGATGAA 1902
Dh 1138 CAAAGGCTGTATTTGCTCGTGTGTTTGAAGAGCGCTCCGCTGATGTTTTCAGAG 1197
Qy 1903 GCTACTTCATGTTAGTTCGATTTACTGAAAGACTATTTCTGTGTCAT-----GAAG 1956
Dh 1198 GCTACAGTGTCTGATACACACAGAGCAGCCTTTCACACATTCACAGAGAC 1257
Qy 1957 GATGTGTCAAACACAGAACTTCTATTTTCATTCACACAGATTTCAACAGTGTGAT 2016
Dh 1258 TTTTCTTCAATTCAAAGACGAGCGTTTACGTTCCCATGACCTGCCACATGCTGAT 1317
Qy 2017 GCAGATGAATCATTTGCTTGGATCAGGGTAAAGTAGCCGAGCGTGAATCCCATGAT 2076
Dh 1318 GCAGATGAATCATTTGCTTGGATCAGGGTAAAGTAGCCGAGCGTGAATCCCATGAT 1377
Qy 2077 TTGCTTCTAACCTCATATGATTTATTCAGAAATGTCAGATACAGAGAGCC 2131
Dh 1378 CTGTTAGCGTCACAAAGATCCCTATACCGGGGCTGTGGGATATTCAGAAAAACC 1432

OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.5%; Score 316.6; DB 21; Length 1185;
Best Local Similarity 58.4%; Pred. No. 1,7e-80;
Matches 594; Conservative 0; Mismatches 414; Indels 9; Gaps 2;

QY 1132 AACTTGGTCMAAGTGCCTATTTCAGTGTGCTTTAAACGCTAATATGTCGCCAGT 1191
DB 1 AACTTGGACAAAGTATCATATTAGTACAGCTTATCAACAGCAATGCTGTGCTCT 60
QY 1192 CAGGAATGTGGCAGGTACCTTACTGTGGAGATCTAGTAATGGTAATGAGCTGCTT 1251
DB 61 CAGGTATCATGAATGGCCAGATGACAGTCGTAATTTGGTATGGTAATGAGCTTCTC 120
QY 1252 TTTGAGCTTTCATTACCCCTGAACTTTCTGGAACTGTATATAGAGACTAGCAAGCA 1311
DB 121 TTTGAGCTGTCTTCCCTTAACTTTTGGTAGTGTTTACCGTAACATTCAGAGC 180
QY 1312 CTCATGATATGAACACCTTGTCTTACTCTACTCAAGTAGACACCCAAATTAAGACAA 1371
DB 181 CTGTGACATGAATCAATCATTTCCAGTGTAGAGAGAAATCTGACATCAAAATPACA 240
QY 1372 GTGATGCAATCTCCCTTCAGATCAACACACAGACGTAACCGGCTTGTATATG 1431
DB 241 AGTAGCAAAAGCCTCTTGTACTGAAGGT-----GAAACATTTGAGTGAAGTGA 294
QY 1432 CATTTGAATACATTGAGGGCCAGAAAGTCTTAGTGAATATCTTTGAAGTCCCTGCA 1491
DB 295 CACTTCAGTTACCTTCCAGAGAGAAATATATATAGTGAATTTCTTTGCTGACCGCA 354
QY 1492 GGAAGAAAGTGCCCATTTGTAGAGGTAGTGGTCAGAGAAAGCACAAATGAGGCTA 1551
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 355 GAAAAAGTGGGCAATCTTGTACTAGCGGAGCGGCAATCAACATCTTAAAGATG 414
 QY 1552 TTAATTCGCTTCTATAGACCTCAAAAAGGATAGATTATCTTGTGCTCAAAATATACAA 1611
 Db 415 CTATTCAGATCTTCTGACACAGATTCGGAATATTAAGATTTGACGGGCAAGATATAAG 474
 QY 1612 GATGTAGCCCTGGAAAGCTTCGAGAGGCAATGGAGTGTACTCTCAGATGCTGTCTC 1611
 Db 475 GAAATGAGACTAGATGAGCTCCGAAAGCTCCATAGAGTTGCGCGCAAGATATCTGTCTC 534
 QY 1672 TTCCATAATCTATTTATTAACACCTTTATATGAAACATCAGTGTCTCACTGAGGAA 1731
 Db 535 TTCAATGACAAATTTTCAACACCTATATATGACGCTCTCACTCACTCAAGAAAGAG 594
 QY 1732 GTGTATGACAGTGGCAAAATTAAGTCTGACTATGATGCAATTTCTGATCCCATGGA 1791
 Db 595 GTGTATGAGGCTGCTCGCGCTGCTGCAATTCATGAAACATCAGTAATTTCCCTGATAA 654
 QY 1792 TATGACACCCAAAGTGGGAAACGAGACTCAAGCTTTGAGAGGAAAGCAAAAGATA 1851
 Db 655 TATTCACAGATCTGTGGGAAAGGGGCTCAAGTTGAGTGTGAGAGAAACAAAGAGTG 714
 QY 1852 GCAATGTGACAGCATTTTAAAGAACCCCACTGATCTATGATGAGCTACTTCA 1911
 Db 715 GCACTGCTCGACATTTTGAATCCCTCGATCTGTGTGTGATGAGGCAACAGAT 774
 QY 1912 TCGTATGATGCTTACTAGAGAGCTATCTTGTGCTGATGAGAGATGCTGCAAAAC 1971
 Db 775 GCGCTGACACAGACACAGAGAGATCTGAAACGACTCAAGGCACTAGCGAGTAC 834
 QY 1972 AGAATCTTATTTTCAATGCAACAGATTTGCAACAGTGTGATGAGATGAAATCAT 2031
 Db 835 AGGACATCTATCTTCAATGCTCCACAGGCTGACCACTGAGATGACAGATGCTGA 894
 QY 2032 GTCTGATGACAGGATGAGTGGCGAAGCTGTACCACTGATGCTGCTGCTGCTGCT 2091
 Db 895 GTACTGGAACCGGAGAGTGTGAAACAGGCCCCACGACGACTATTTGGGAAATC- 953
 QY 2092 CATGATCTATTTGAGAAATGTGCAATACAGAGACGCTGTGCAAGAACATGAT 2148
 Db 954 -AAGGAGATATGCGCAGCTATGACACAAACAGCAGCGTGTGATTCCTGAT 1008

RESULT 14
 ABL28787 standard; DNA; 2770 BP.
 ID ABL28787
 AC ABL28787;
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 37834.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN NO200171042-A2.
 XX
 PD 27-MAR-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX WPI; 2001-656860/75.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 37834; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB101849), expressed DNA
 CC sequences (AB101840-AB101849), and the encoded proteins
 CC (AB101840-AB101849).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2770 BP; 629 A; 713 C; 725 G; 703 T; 0 other;
 Query Match 12.8%; Score 299.6; DB 23; Length 2770;
 Best Local Similarity 50.7%; Pred. No. 2.2e-75;
 Matches 778; Conservative 0; Mismatches 749; Indels 9; Gaps 2;
 QY 585 CTATGTGTATCAAGAGCTGAGCTGCTTTTAAACGAATGCAATGCAATTTGG 644
 Db 1063 CCAAGGTTGGGCGGACATGTGATGATGATTTATTAATCTTGTGCTCTGAT 1122
 QY 645 CAGGTGAGCCCAATTTCAATCCGAAATATGCCAAATATGCTTTCTCATCTTCA 704
 Db 1123 ACAGGTTCAACAGTACACGACCGGAGATGAGATTGACCTGTCCGATCTGACCA 1182
 QY 705 CTGTGATCTGGTTTTCACCTGAGACAGACGAGGCTTTATCTAAGCTATTGACAG 764
 Db 1183 GCTTTCCTTATGATGAGACTTGCACGCAAAACGAGAGAGTGTGCTCATGATCG 1242
 QY 765 AGAACAAGGGGATATCAATTTTCTGAGTCTTGTGATTTAATCTTCTCCATCAT 824
 Db 1243 CGGACACGAGCTGATCAACCACTCTCACTATTTGCTTCTGATCCACCAACCAT 1302
 QY 825 GTTGAAGTATGCTTCTGAGTGTGTTTGTATTAATGCGGCTCCAGTTGCTT 884
 Db 1303 TCTGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
 QY 885 GGTAACTCTTGAACACTTGTGATACACAGATTTCACTGAGTCACTGAGTGGAG 944
 Db 1363 AATGCTTCTCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1422
 QY 945 AACTAGATTTGAATGAATGAATGAACAAAGCAGATTAATGATGATGCTGCTATGA 1004
 Db 1423 AACCAAGTACCAAGAGGAGATGAACCTGGGACATGAGCAGAGCTGGCAGCTGGA 1482
 QY 1005 CTGACTGCTGAATTTGAATCTGTAATTTTAATTAATGAAGATATGAAGACAGAG 1064
 Db 1483 CTCCTTCTCAATTTGGAACAGTCAAGTCAAGGACGACACATCAAGGTGAGATTG 1542
 QY 1065 ATATGATGATTTTGAAGCTATGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1124
 Db 1543 CTACAGGAGGCAATCTCAAGTACCAAAAGAGAGTCTCTCAATGCTGACACTCAA 1602
 QY 1125 TATGCTGAATTTGTGCAAGTCTATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 1184
 Db 1603 CATGCTGAACACAGCCAGAAATATTTTGTGCTGGGCTGCTGCTGCTGCTGCTGCT 1662
 QY 1185 C---GCCAGTCAAGGAATTTGCGAGTACCTTACTGTTGAGATCTAATATGTTGAA 1241
 Db 1663 CTGTGTATCTTGTGATGATCAATCAAAACCTCACCGTGGGAGATTGTTGCTTTCTC 1722
 QY 1242 TGAGCTGCTTTTCAAGCTTCAATTAACCCGTAATTTTGGGAACTGTATATGAGAGAC 1301
 Db 1723 CACTTATCTGATGAAATTTGATCAATGCACTGAATTTGCTGGCACTTATTAACGAGCAT 1782
 QY 1302 TAGACAAGCACTCATGATATGAACAATTGTTACTTACTCACTCAAGGTAGACCCCAAT 1361


```

Db 1783 ACAGAGAACTTCGTGACATGAGAGAACTGTCATCTGTGAAAGAGAGAGAT 1842
Oy 1362 TAAAGCAAAAGATGATGATCTCCCTTCAGATCAACACAGAGACATACCGTGGCTT 1421
Db 1843 TGTGATGCCCCAGGCTGTTCCCTCTTAACTGCC-----GGCGGGGATTTAGATT 1896
Oy 1442 TGTATATGTCATTTTGAATATACATTGAGGGCCAGAAAGTCCTTAGTGAATATCTTTGA 1481
Db 1897 CTCAAACGTGACCTTTGGCTATTACCCGAGAGAAATAGTCTGCCAATGTAGCTTAC 1956
Oy 1482 AGTCCCTGAGAGAAAGAAAGTGGCCATTGTAGAGGTAGTGGTACGAGAAAGACAAAT 1541
Db 1957 AGTTCGGCGGTAGAGACGGTAGCATGTTGGACCTCTCGAGCGGGTAAGACACAT 2016
Oy 1542 AGTGAAGCTATATTTCCGTTCTATGAGCCTCAAAAGGTACATTTATCTTCTGTGCA 1601
Db 2017 CATGGCGGCTGCTTTCCGCTTTTACGATGTGACACTGCTATCTGATGACGGACA 2076
Oy 1602 AATATATCAAGATGTAGAGCTGAGAAAGCTTGGAGGGAGTGGAGTGGTACTCAGGA 1661
Db 2077 AATATCAAGCTTGTGACGACAGAGTTCGAAAGGCCATTGAGTGGTCCCGCAGGA 2136
Oy 1662 TGCTGTCTCTTCCATATATATATATATACAACTCTTATATGAAACATGAGCTTC 1721
Db 2137 TACGGTCTCTTAAACAACACCATCTTCAACAATGAAATGCAAGTGGGCGCTTC 2196
Oy 1722 ACCTGAGAGAGTGTATGACGTGGCAAAATTAAGTGAATGATGATGATCTTCTGAAAT 1781
Db 2197 GGACGAGAGGTGTATGAGGCTGCTCGCGCCCGCAGCATTCAGAAAGATTTCTGGCTT 2256
Oy 1782 GCCACATGATATGACACCCAGTAGAGGAGAGAGACTCAAGCTTTTCAGAGAGAAAA 1841
Db 2257 TCCGAGAAATATGAGACAAAGTGGGAGACAGAGTCTCGGCTCAGTGGAGGTGAAA 2316
Oy 1842 GCAAGAGTAGCAATTTGCAAGAGCCATTTTGAAGAACCCCGCAGTCAATCTATGATGA 1901
Db 2317 GCAAGCGGTAGCATGCGAGAGACCTTCTTAAGGGGCCATTAATGTGCTCTGGATGA 2376
Oy 1902 AGCTACTCATGCTAGATTCATTAAGAGACTATCTTGTGGCCATGAAGAGAT 1961
Db 2377 AGCCACCTGAGTTTGGAGACCCACAGACGAAATATTCAGGCTTTGGCCAGGCT 2436
Oy 1962 GGTCAACACAGAACTTATTTTATTCAGACAGATGTTCAAGAGTGTGATGACAGA 2021
Db 2437 GTGCCAAATCTACACCATATGTTGCCATGCTCTTCCACATCATCATACGCGA 2496
Oy 2022 TGAATTCATTTGTTGGATCAGGGTAAAGTACCCGAACTGTGTACCCCATGTTTGT 2081
Db 2497 CGAGATTTCTGTCTGACAGCAGGAAAGCATACGAGCGTGGCGCACAGAGAGCTTGT 2556
Oy 2082 TGTCTAACCTCATATGATCTATTCAGAAATGTGGCA 2117
Db 2557 GCTGGCGAGAGATGCATCTACGCGGACATGTGGCA 2592

```

RESULT 15

AAD27254
ID AAD27254 standard; cDNA, 2984 BP.

AC AAD27254;

DT 09-APR-2002 (first entry)

XX Human transporter and ion channel-1 (TRICH-1) cDNA.

XX Human; transporter and ion channel-1; TRICH-1; neuroprotective; asthma;
KW noctropic; cytosolic; cardiovascular; immunosuppressive; cardiomyopathy;
KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukemia;
KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;
KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;
KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;
KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;

```

KW bradyarrhythmia; gene expression; drug screening; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 269..2797
FT /*tag= a
FT CDS /product= "Human TRICH-1 protein"
PN WO200192304-A2.
PD
PD 06-DEC-2001.
PF
PF 25-MAY-2001; 2001MO-US17065.
XX
PR 26-MAY-2000; 2000US-208424P.
PR 01-JUN-2000; 2000US-209001P.
PR 08-JUN-2000; 2000US-210588P.
PR 16-JUN-2000; 2000US-212335P.
PR 22-JUN-2000; 2000US-213747P.
PR 29-JUN-2000; 2000US-215391P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Thornton M, Walla NK, Yue H, Nguyen DB, Lai P, Gandhi AR;
XX Triboley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YT;
XX Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjambala MS;
XX Raumann BE, Lee EA, Hafalia A, Baughn MR, Green BD, Khan FA;
XX Kearney L, Elliot VS, Seilhammer JU, Policky JU, Borowsky ML;
XX Burford N, Ding L, Lu DM, Hallman JL;
XX
XX WPI; 2002-122055/16.
XX P-PSDB; AAE16764.
XX
XX New human transporters and ion channels (TRICH) polypeptides useful for
XX diagnosing, treating or preventing disorders associated with aberrant
XX expression of TRICH
XX
XX Claim 11; Page 187-188; 210pp; English.
XX
XX The invention relates to human transporters and ion channels (TRICH)
XX polypeptides and their cDNA molecules. The nucleic acid and polypeptide
XX sequences are useful in the diagnosis, treatment, and prevention of
XX disorders associated with transport (akinesia, cystic fibrosis, Bell's
XX palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,
XX amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's
XX muscular dystrophy); immunological (AIDS, Addison's disease, allergies,
XX asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);
XX cardiac disease (angina, hypertension, or bradyarrhythmia) and in the
XX assessment of the effects of exogenous compounds on the expression of
XX nucleic acid and amino acid sequences of transporters and ion channels.
XX The polynucleotides may be used to detect and quantify gene expression
XX in biopsied tissues in which TRICH expression may be correlated with a
XX disease, to generate hybridization probes for mapping naturally occurring
XX genomic sequence, and in drug screening. The present sequence is human
XX TRICH-1 cDNA.
XX
XX Sequence 2984 BP; 571 A; 831 C; 916 G; 666 T; 0 other;
XX
XX Query Match 12.1%; Score 284.2; DB 24; Length 2984;
XX Best Local Similarity 50.6%; Pred. No. 6,4e-71;
XX Matches 744; Conservative 0; Mismatches 718; Indels 9; Gaps 2;

```

Db 1434 TCAATGTCATCCCAAGCTGGCCGACATCATCTTGGCATCTACTTGCAGATGTTCT 1493
 QY 866 GCGGTGCCAGTTGCTTTGGTAACTTGGAAACCTTGGAAACCTTGGTATACACAGCATTTCAAG 925
 Db 1494 TCAAGCGCTGGTGGTTCATGTTGTTCTGTGATGAGTCTTACTGACCTCGAACCA 1553
 QY 926 TTGCACTACACAGGTGGGAACTAGATTAGAAATGAAATGAAACAGAGCATTAATGATG 985
 Db 1554 TTGGGTCTACGAGTGGGAAACCAAGTTTCGTGCTATTAACACACAGAGAAACGCTA 1613
 QY 986 CAGGTAATGCTGCTATAGACTCACTGCTGAATTAATGAACCTGGAAGTATTTAATATG 1045
 Db 1614 CCCGGGACGAGCACTGAGTACTCTTGCTAACTTGGAACTGGAACGTTGAATTAACAACCGC 1673
 QY 1046 AAGATATGAAGCAACAGATATGATGATTTTGAACCTGATGAGACTGCTTCAATGA 1105
 Db 1674 AGAGTTACGAAGTGAACGCTATCGAAGGCCATCATCAATATCAAGGTTTGAAGTGA 1733
 QY 1106 AAAGTACTCTACTCTGCTATGCTGAACCTTGGTCAAAAGTCTAATTTTCACTGCTG 1165
 Db 1734 AGTGAAGCGCTTCACTGTTTACTTAATCAGACCCAGAACCTGATGAGGCTCGGCG 1793
 QY 1166 TAAAGCTATATATGCTGCTGCGCAGTCAAGGAAATTTGGGAGTACCTTACTGTTGAG 1225
 Db 1794 TCCGCGCGGCTCCCTGCTTTGGCGATATCTTGTACTGAGCAGAGACTACAGGTTGAG 1853
 QY 1226 ATCTAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1285
 Db 1854 ACTATGCTCTTGGGACCTACATTTTCACTGATGATGATGATGATGATGATGATGATG 1913
 QY 1286 CTGATATATGAGAGACTAGACAGACACTCATATATGAAACCTTGTACTTACTCA 1345
 Db 1914 CTTACTACAGATATATCAGACCAACTTCAATGACATGAGAACATGTTGACTTGTCTGA 1973
 QY 1346 AGGTAGACACCAATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATG 1405
 Db 1974 AAGGAGAGACGAAGTGAAGGACCTTCTGAGAGAGGCGCTTCTGCTT-----CAGA 2027
 QY 1406 CAGCTACCTGCTGCTTGAATATGATGATGATGATGATGATGATGATGATGATGATG 1465
 Db 2028 AGGCGCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2087
 QY 1466 GTGGAATATCTTTGAATGCTGCTGAGGAAAGATGCTGATGATGATGATGATGATG 1525
 Db 2088 AGGACGCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2147
 QY 1526 CAGGGAAGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1585
 Db 2148 CAGGGAAGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2207
 QY 1586 TTTATCTTGTGCTCAAAATATCAAGATGATGATGATGATGATGATGATGATGATGATG 1645
 Db 2208 TCCGAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2267
 QY 1646 GAGTGTACTCTGAGATGCTGCTCTTCAATTAATTAATTAATTAATTAATTAATTAATG 1705
 Db 2268 GAGTGTGCTGCTGAGATGCTGCTCTTCAATTAATTAATTAATTAATTAATTAATTAATG 2267
 QY 1706 GAAACATCAAGTCTTCACTGAGAGATGATGATGATGATGATGATGATGATGATGATG 1765
 Db 2328 GCCGTGTCAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2387
 QY 1766 ATGCAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1825
 Db 2388 ATGCAATTTATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2447
 QY 1826 TTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1885
 Db 2448 TGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507
 QY 1886 TCAATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1945
 Db 2508 TCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2567

QY 1946 GTGCCATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2005
 Db 2568 CTTCCTGCGCAAAAGTGTGTCACACCGACACATCGATGATGATGATGATGATGATGATGATG 2627
 QY 2006 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2065
 Db 2628 CTGTGTCAATGCTGACAGATCTCTGATCAAGATGATGATGATGATGATGATGATGATGATG 2687
 QY 2066 CCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2125
 Db 2688 GACACGAGGCTCTGTTGTC--CGAGTGGGCTGATGCTGATGATGATGATGATGATGATG 2744
 QY 2126 GCAGCGGTGTGAGAAACATGATTAACCCCA 2156
 Db 2745 AGGGAACAGAAAGAACTCTGAAAGACATTA 2775

Search completed: February 15, 2003, 01:40:17
 Job time : 512 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 01:30:25 ; Search time 87 Seconds
(without alignments)
8266.172 Million cell updates/sec

Title: AF133659
Perfect score: 2345
Sequence: 1 ATGCGCGCTCTCGCATGCA.....TATTAAAAAATCATCACTT 2345

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A COMB. seq.*
2: /cgn2_6/prodata/2/ina/5B COMB. seq.*
3: /cgn2_6/prodata/2/ina/6A COMB. seq.*
4: /cgn2_6/prodata/2/ina/6B COMB. seq.*
5: /cgn2_6/prodata/2/ina/PCBUS COMB. seq.*
6: /cgn2_6/prodata/2/ina/bcktlf1es1. seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2248	95.9	2407	2	US-08-895-522-2
2	2248	95.9	2407	3	US-09-195-591-2
3	358.6	15.3	5369	4	US-09-212-247C-3
4	188.8	8.1	1749	4	US-09-134-001C-1893
5	180.2	7.7	7760	4	US-08-961-527-63
6	170	7.2	5120	3	US-08-772-370A-6
7	170	7.2	8370	2	US-08-488-706-1
8	165.8	7.1	13188	4	US-08-961-527-70
9	165.4	7.1	3576	4	US-08-976-259-79
10	154.8	6.6	6492	4	US-08-961-527-188
11	147	6.3	7147	4	US-08-961-527-23
12	143.8	6.1	2726	1	US-08-461-823-1
13	143.8	6.1	4264	2	US-08-784-649A-1
14	143.8	6.1	4264	2	US-08-784-649A-5
15	143.8	6.1	4646	1	US-08-181-471-2
16	143.8	6.1	4659	6	5206352-3
17	143.8	6.1	6505	2	US-08-793-610-5
18	143.8	6.1	9318	2	US-08-793-610-6
19	142.2	6.1	4659	2	US-08-583-276-18
20	140.6	6.0	4659	2	US-08-752-447-1
21	139.4	5.9	1201	3	US-09-316-167-1
22	139.4	5.9	1201	3	US-08-961-083-189
23	139.4	5.9	7186	4	US-08-961-527-39
24	135.2	5.8	1743	4	US-09-134-001C-2774
25	134.8	5.7	11864	4	US-08-961-527-61
26	131	5.6	7721	3	US-08-772-370A-14
27	129.4	5.5	4233	3	US-09-120-513-1

28	129.4	5.5	4233	4	US-09-450-105-1	Sequence 1, Appl1
29	126	5.4	775	4	US-09-221-017B-860	Sequence 860, App
30	125.8	5.4	4002	2	US-08-996-545-1	Sequence 1, Appl1
31	125.8	5.4	4002	2	US-08-996-545-3	Sequence 3, Appl1
32	125.8	5.4	4002	4	US-09-328-320-1	Sequence 1, Appl1
33	125.8	5.4	4002	4	US-09-328-320-3	Sequence 3, Appl1
34	125.2	5.3	4047	2	US-08-612-734B-1	Sequence 1, Appl1
35	125.2	5.3	4800	2	US-08-612-734B-3	Sequence 3, Appl1
36	116.6	5.0	28473	4	US-08-961-527-83	Sequence 83, Appl1
37	116.2	5.0	4224	1	US-08-612-521-1	Sequence 1, Appl1
38	112.6	4.8	1749	4	US-09-221-017B-703	Sequence 703, App
39	109	4.6	14672	4	US-08-961-527-111	Sequence 111, App
40	108.8	4.6	5595	1	US-07-841-655-1	Sequence 1, Appl1
41	108.8	4.6	5595	1	US-07-635-965C-1	Sequence 1, Appl1
42	108.6	4.6	3909	1	US-08-232-537-1	Sequence 1, Appl1
43	107.8	4.6	1971	4	US-08-858-207A-72	Sequence 72, Appl1
44	107.8	4.6	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
45	107.6	4.6	3924	2	US-08-996-644-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-895-522-2
; Sequence 2, Application US/08895522
; Patent No. 5858719
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purni
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,522
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0336 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2407 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: OVARNOT02
; CLONE: 545981
; US-08-895-522-2
Query Match 95.9%; Score 2248; DB 2; Length 2407;

Best Local Similarity 99.4%; Pred. No. 0;
Matches 2265; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

OY	8	TGTCGCGATTCATCTTCTTGCGCGCTCGCGCGCGCGCGCGCTGCTTTCCGAAAAGCGCCGCGC	67
Db	1	TGTCGCGATTCATCTTCTTGCGCGCTCGCGCGCGCGCGCGCTGCTTTCCGAAAAGCGCCGCGC	60
OY	68	ACTCCGCAATTCATCCGGCCTTTAGTCTCTGTAGCGGCTCAGGTCGCGAGTGAAGGC	127
Db	61	ACTCCGCAATTCATCCGGCCTTTAGTCTCTGTAGCGGCTCAGGTCGCGAGTGAAGGC	120
OY	128	CACATCAACTCGCGCCTTGCGAAACGCGCTCGAGCTTACAGATTCGAGTCATTPAAAA	187
Db	121	CACATCAACTCGCGCCTTGCGAAACGCGCTCGAGCTTACAGATTCGAGTCATTPAAAA	180
OY	188	GTAATCACTGGCAGAAATTTGGGAAAAGGCAATTCAGACAGTTCTTAAGATGCTGGAAAAG	247
Db	181	GTAATCACTGGCAGAAATTTGGGAAAAGGCAATTCAGACAGTTCTTAAGATGCTGGAAAAG	240
OY	248	CTCTCCAGTATGCGCCACTGTATGAAAAGAGACATGTTGGCATGGTCAATGCAGAGAG	307
Db	241	CTCTCCAGTATGCGCCACTGTATGAAAAGAGACATGTTGGCATGGTCAATGCAGAGAG	300
OY	308	GACTCCACACAGACCCAAAAGAGGTTAAAGATGTGTACTCTCGAAAATCATPAAAG	367
Db	301	GACTCCACACAGACCCAAAAGAGGTTAAAGATGTGTACTCTCGAAAATCATPAAAG	360
OY	368	CAATGCTTTCTTATGTGTGCGCCAAAACAGCGCCATCTACGAGCTAAGTTGGCATTT	427
Db	361	CAAAAGCTTTCTTATGTGTGCGCCAAAACAGCGCCAGATCTACGAGCTAAGTTGGCATTT	420
OY	428	CGCTGGGATTTTTGGGTGTGTCAAAAGGCCATGAAATATGTGTGTTCCCTTCATGTTAAAT	487
Db	421	CGCTGGGATTTTTGGGTGTGTCAAAAGGCCATGAAATATGTGTGTTCCCTTCATGTTAAAT	480
OY	488	ATGCTGTAGACAGCTTCAACAGATGTGCGGAAAACATGCTGAACCTGAGTATGCACCA	547
Db	481	ATGCTGTAGACAGCTTCAACAGATGTGCGGAAAACATGCTGAACCTGAGTATGCACCA	540
OY	548	ATACAGTTGCAACCATGSCAACAGCAATTCGTAATGGCTATATGCTGTATCAAGAGCGGAG	607
Db	541	ATACAGTTGCAACCATGSCAACAGCAATTCGTAATGGCTATATGCTGTATCAAGAGCGGAG	600
OY	608	CTGCTTTTTTAAAGAAATTCGAATCAGTATTTGGCAGGTAGCCCAAAATTCATCC	667
Db	601	CTGCTTTTTTAAAGAAATTCGAATCAGTATTTGGCAGGTAGCCCAAAATTCATCC	660
OY	668	GAAGATATGCCAAAATATGCTTTCTGCATCTTCACAAACCTGATCTGGGTTTTCACTGA	727
Db	661	GAAGATATGCCAAAATATGCTTTCTGCATCTTCACAAACCTGATCTGGGTTTTCACTGA	720
OY	728	GCAAGCACAAGGAGCTTTATCTTAAGGCTATTCAGAGAGAAACAAGGGTATCAGTTTTG	787
Db	721	GCAAGCACAAGGAGCTTTATCTTAAGGCTATTCAGAGAGAAACAAGGGTATCAGTTTTG	780
OY	788	TCTCTAGGCTTTGGTATTTATCTTCTTCCCATCATGTTTGAAGTATGCTTGTCACTG	847
Db	781	TCTCTAGGCTTTGGTATTTATCTTCTTCCCATCATGTTTGAAGTATGCTTGTCACTG	840
OY	848	GTGTTTGTATTAACAATGCGGTGCCAGATTGCTTGGTAAACCTTGGAAACATTGGTA	907
Db	841	GTGTTTGTATTAACAATGCGGTGCCAGATTGCTTGGTAAACCTTGGAAACATTGGTA	900
OY	908	CATACACAGCAATTCACAGTTGCATCAACAGGTGGAGAACTAATATTAAATPAGAAATGA	967
Db	901	CATACACAGCAATTCACAGTTGCATCAACAGGTGGAGAACTAATATTAAATPAGAAATGA	960
OY	968	ACAAGAAGATATGATGACAGGTATATCTGTATAGACTCACTGCTGAATTTATGAACCTG	1027
Db	961	ACAAGAAGATATGATGACAGGTATATCTGTATAGACTCACTGCTGAATTTATGAACCTG	1020
OY	1028	TGAAGTATTTTAAATATGAAGAATATGAAACACACAGATATGATGATTTTGAAGACGT	1087

Db	1021	TGAAGTATTTTAATATATGAAGAATATGAAGCAGAGATATATGATGATTTTGTGAAGAGCT	1088
Qy	1088	ATGAGACTGCTTCATTTGAAAAAGTACTCTACTCTGGCTATGCTGAACTTTGGTCAAAGTG	1147
Db	1081	ATGAGACTGCTTCATTTGAAAAAGTACTCTACTCTGGCTATGCTGAACTTTGGTCAAAGTG	1140
Qy	1148	CTATTTTCAAGTGCGGTTTAAACAGCTAATAATGGTCTCCGACAGTCAGGAAATTGGGAG	1207
Db	1141	CTATTTTCAAGTGCGGTTTAAACAGCTAATAATGGTCTCCGACAGTCAGGAAATTGGGAG	1200
Qy	1208	GTAACCTTACTGTGGAGATCTAGTAATGGTAATGAGTACTGTTTTTACGCTTTCATTAAC	1267
Db	1201	GTAACCTTACTGTGGAGATCTAGTAATGGTAATGAGTACTGTTTTTACGCTTTCATTAAC	1260
Qy	1268	CCCTGAACCTTCTGGGAACTGTATATAGAGAGACTGACAAAGCTCATAGATATGAACA	1327
Db	1261	CCCTGAACCTTCTGGGAACTGTATATAGAGAGACTGACAAAGCTCATAGATATGAACA	1320
Qy	1328	CCTTGTTTACTCTACTCAAGGTAGACACCCCAATTAAAGACAAAGTGATGCTCTCCC	1387
Db	1321	CCTTGTTTACTCTACTCAAGGTAGACACCCCAATTAAAGACAAAGTGATGCTCTCCC	1380
Qy	1388	TTCAATACACACACAGACAGCTACCGTGGCCTTTGATATGTGATTTTGATATCATTTG	1447
Db	1381	TTCAATACACACACAGACAGCTACCGTGGCCTTTGATATGTGATTTTGATATCATTTG	1440
Qy	1448	AGGGCAGAAAGTCTTATGTGGAAATATCTTTTGAAGTCCCTGCAGAGAAAGAGTGCCA	1507
Db	1441	AGGGCAGAAAGTCTTATGTGGAAATATCTTTTGAAGTCCCTGCAGAGAAAGAGTGCCA	1500
Qy	1508	TTTGAGAGATATGTGGGTCAAGGAAAGACACAAATGTATAGGCTATTAATTTGGCTCTATG	1567
Db	1501	TTTGAGAGATATGTGGGTCAAGGAAAGACACAAATGTATAGGCTATTAATTTGGCTCTATG	1560
Qy	1568	AGCCTCAAAAGGGTATGACTTTATCTTGGCTGCAAAATATCAAGATGTGACCTGAGAA	1627
Db	1561	AGCCTCAAAAGGGTATGACTTTATCTTGGCTGCAAAATATCAAGATGTGACCTGAGAA	1620
Qy	1628	GCCTTCGAGGGCAGTGGGAGTGTGACTCAGAGATGTGTCTCTTCATATACTATTT	1687
Db	1621	GCCTTCGAGGGCAGTGGGAGTGTGACTCAGAGATGTGTCTCTTCATATACTATTT	1680
Qy	1688	ATTACAACTCTTATATGGAACAATCAAGTCTTCACTGAGAGAAAGTATATGCAATGGCA	1747
Db	1681	ATTACAACTCTTATATGGAACAATCAAGTCTTCACTGAGAGAAAGTATATGCAATGGCA	1740
Qy	1748	AATTAGCTGGAATTCATATGCAATCTTCGAAATGCCACATGAGATATACACCCCAATG	1807
Db	1741	AATTAGCTGGAATTCATATGCAATCTTCGAAATGCCACATGAGATATACACCCCAATG	1800
Qy	1808	GGGAACGAGAGACTCAAGCTTTCAGAGAGAGAAAGCAAAAGATGACATTCGAGAGCCA	1867
Db	1801	GGGAACGAGAGACTCAAGCTTTCAGAGAGAGAAAGCAAAAGATGACATTCGAGAGCCA	1860
Qy	1868	TTTTGAAGGACCCCCAGTCACTATATGATGAAGCTACTTCATCGTTAGATTGATTA	1927
Db	1861	TTTTGAAGGACCCCCAGTCACTATATGATGAAGCTACTTCATCGTTAGATTGATTA	1920
Qy	1928	CTGAAGAGACTATCTTGCTGTCAGAAAGATGTGGTAAACACAGAACTCTATTTTCA	1987
Db	1921	CTGAAGAGACTATCTTGCTGTCAGAAAGATGTGGTAAACACAGAACTCTATTTTCA	1980
Qy	1988	TTGCAACAGATTTTCAACAGTGTGTGATGCAAGATGAATCATGTCTTGGATCAGGTA	2047
Db	1981	TTGCAACAGATTTTCAACAGTGTGTGATGCAAGATGAATCATGTCTTGGATCAGGTA	2040
Qy	2048	AGGTAGCCGAACGTGGTACCAACATGGTTTGTCTTAAACCTCATATGATCTATTTAG	2107
Db	2041	AGGTAGCCGAACGTGGTACCAACATGGTTTGTCTTAAACCTCATATGATCTATTTAG	2100
Qy	2108	AAATGTGGCATACACAGACAGCCGTGTGCGAACAATGATTAACCCCAATGGGAAGCA	2167
Db	2101	AAATGTGGCATACACAGACAGCCGTGTGCGAACAATGATTAACCCCAATGGGAAGCA	2160

QY	1268	CCGGAACCTTTCGTGGAAACGTGATATATGAGACCTGACAGACGACTGATAGATATGACA	1327
Dd	1261	CCCTGAACCTTTCGTGGAACTGTATATGAGAGCTGACAGACCTCATAGATATGAGACA	1320
QY	1328	CTTGTGTACTCTACTCAGGTAGACACCCAAATTAAAGCAAAAGTATGGGACTCTCCC	1387
Dd	1321	CTTGTGTACTCTACTCAGGTAGACACCCAAATTAAAGCAAAAGTATGGGACTCTCCC	1380
QY	1388	TTTCAGATCACACACACAGACGCTACCGTGGCCCTTTGATATGTGCAATTTTGAATACATTG	1447
Dd	1391	TTTCAGATCACACACACAGACGCTACCGTGGCCCTTTGATATGTGCAATTTTGAATACATTG	1440
QY	1448	AGGGCCAGAAAGTCCCTTAGTGGAAATATCCCTTGAAGTCCCTGACGAGAAAGAAAGTGCCA	1507
Dd	1441	AGGGCCAGAAAGTCCCTTAGTGGAAATATCCCTTGAAGTCCCTGACGAGAAAGAAAGTGCCA	1500
QY	1508	TTGTAGAGAGTAGTGGGTACAGGGAAAAGCACAATAGTAGGACTATTTTGGCTTCTATG	1567
Dd	1501	TTGTGTAGAGAGTAGTGGGTACAGGGAAAAGCACAATAGTAGGACTATTTTGGCTTCTATG	1560
QY	1568	AGCCCTCAAAAGGGTATGCAATTTATCTTGTCTGTCAAAATATACAAAGATGTGACCTGGAAA	1627
Dd	1561	AGCCCTCAAAAGGGTATGCAATTTATCTTGTCTGTCAAAATATACAAAGATGTGACCTGGAAA	1620
QY	1628	GCCCTTCGAGGGGACAGTGGGAGTGGTACCTCAGAGATGCTGCTCTCCATATACATATTT	1687
Dd	1621	GCCCTTCGAGGGGACAGTGGGAGTGGTACCTCAGAGATGCTGCTCTCCATATACATATTT	1680
QY	1688	ATTACACCTCTTATATATGAAAACATCAGTGCCTTCACTGAGAAAGTATGCAATGGAGCA	1747
Dd	1681	ATTACACCTCTTATATATGAAAACATCAGTGCCTTCACTGAGAAAGTATGCAATGGAGCA	1740
QY	1748	AATTAGCTGGACTTCATATGATGCAATTTCTTGAAATGCCATGGAATATGACACCCAAAGTATG	1807
Dd	1741	AATTAGCTGGACTTCATATGATGCAATTTCTTGAAATGCCATGGAATATGACACCCAAAGTATG	1800
QY	1808	GGGAAACGAGGACTCAGCTTTCAGAGAGAGAAAGAAAGAGTAGAGCAATTTGCANAAGCA	1867
Dd	1801	GGGAAACGAGGACTCAGCTTTCAGAGAGAGAAAGAAAGAGTAGAGCAATTTGCANAAGCA	1860
QY	1868	TTTTGAAAGAACCCCCAGCTACTACTATATGATGAGTACTTATCTGTTAAGTGGATTA	1927
Dd	1861	TTTTGAAAGAACCCCCAGCTACTACTATATGATGAGTACTTATCTGTTAAGTGGATTA	1920
QY	1928	CTGAAGACACTATTCTTGATGATGATGAGAGATGAGTCAAAACACAGAACTTCTATTTTCA	1987
Dd	1921	CTGAAGACACTATTCTTGATGATGATGAGAGATGAGTCAAAACACAGAACTTCTATTTTCA	1980
QY	1988	TTTGACACACAGATTGTCAACAGTGGTTGATGCAATGAAATCATTTGCTTGATCAGGGTA	2047
Dd	1981	TTTGACACACAGATTGTCAACAGTGGTTGATGCAATGAAATCATTTGCTTGATCAGGGTA	2040
QY	2048	AGGTAGCCGAAACGTGTATCCCAACATGGTTTGTCTGTAAACCTCTCATATATCTATTCAG	2107
Dd	2041	AGGTAGCCGAAACGTGTATCCCAACATGGTTTGTCTGTAAACCTCTCATATATCTATTCAG	2100
QY	2108	AAATGTGCATACACAGAGACGCGGTGTGACAGAACATGTAACCCCAATGGAGAGCA	2167
Dd	2101	AAATGTGCATACACAGAGAGCGCGGTGTGACAGAACATGTAACCCCAATGGAGAGCA	2160
QY	2168	AGAAAGAAATATATCCAAAGAGAGAGAAAGAAAGAACTACAAAGAAATTTGTCAATA	2227
Dd	2161	AGAAAGAAATATATCCAAAGAGAGAGAAAGAAAGAACTACAAAGAAATTTGTCAATA	2220
QY	2228	GTGTGAAAGGCTGTGTGAAACGTGTCTGTGTCTAATGCTACATAGACATTTCTTTTGTGT	2286
Dd	2221	GTGTGAAAGGCTGTGTGAAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2278

RESULT 3
US-09-212-247C-3
; Sequence 3, Application US/09212247C

```

1 Patent No.6391603
2 GENERAL INFORMATION:
3 APPLICANT: POMPEJUS, Markus; SUEBBERGER, Harald; JOSEFCKEN, Hans
4 MOISGANG, DOVAU, Jose Luis Reveleta; JIMENEZ, Alberto;
5 and GARCIA, Maria Angeles Santos
6 TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii
7 and the use thereof in microbial riboflavin
8 synthesis
9 NUMBER OF SEQUENCES: 21
10 CORRESPONDENCE ADDRESS:
11 ADDRESS: Keil & Weinkauff
12 STREET: 1101 Connecticut Avenue
13 CITY: Washington
14 STATE: D.C.
15 COUNTRY: USA
16 ZIP: 20036
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
19 COMPUTER: IBM AT-compatible, Pentium processor
20 OPERATING SYSTEM: Windows 95
21 SOFTWARE: Wordperfect version 6.1
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/09/212,247C
24 FILING DATE: 16-Dec-1998
25 CLASSIFICATION: <Unknown>
26 INFORMATION FOR SEQ. ID NO: 3:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 5369 base pairs
29 TYPE: nucleic acid
30 STRANDEDNESS: single
31 TOPOLOGY: linear
32 MOLECULE TYPE: DNA (genomic)
33 HYPOTHETICAL: NO
34 ANTI-SENSE: NO
35 FEATURE:
36 NAME/KEY: 5'UTR
37 LOCATION: 1..54
38 FEATURE:
39 NAME/KEY: CDS
40 LOCATION: 55..1482
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 1767..3299
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: 3588..4703
47 FEATURE:
48 NAME/KEY: 3'UTR
49 LOCATION: 4704..5369
50 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
51 US-09-212-247C-3
52
53 Query Match 15.3%; Score 358.6; DB 4; Length 5369;
54 Best Local Similarity 54.1%; Pred. No. 2.6e-94;
55 Matches 777; Conservative 0; Mismatches 649; Indels 9; Gaps 2.
56
57 703 AACCTGATCTGGGTTTTCACCTGACAGACAGACGAGCTTTATCTPAAGGCTATTGAC 762
58 1 AAGCTTGACCTTGCTGGACACTTGAGTCGGCAGACAGAGTGACCTAACCCAGCAATGAT 60
59
60 763 AGAGGAACAAGGGGATACAGTTTGTCTCGTAGCTTTGATTAAATCTTCCCATC 822
61 CCTGGTGTGAAGGATCTCTTATGTGCTCACTGCAATGGTTTTCACATAATATACGATT 120
62
63 823 AAGTTGAAGTATGCTGTGATGAGTGTTTGTATTAACAATGCGGTCAGTTTGTCT 882
64 121 AATTTGAATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
65
66 883 TTGCTAACCTTGGAAACCTTGATACATACACAGCATTCACAGTTTCACAGTCCAGTCA 942
67 181 GGTATTAACATTTGACATATGCTTCTTTACTGTCATCTTTACTTTGAGAACAGACGGCGTGG 240
68
69 943 AGAAGTGAATTTGAATGAATGAACAAGACATTAATGATGACAGGTATGCTCTCTATA 1002

```


Db 241 CGACACGCTTACGCGTGAATGCGCAAGGCTGCATTAAGGCGGCTAGTGTGCATTTG 300
Qy 1003 GACTCAGCTGTAATTAAGAACTGTGAAGTATTTTAATTAAGAAAGATATGAACAG 1062
Db 301 GATTCCCTTAATTAATTTGAAGCTGTAAAGTATTTCAATTAAGAGAACTACCTGGAC 360
Qy 1063 AGATATGATGATTTTGAAGACGTATGAGACTGCTTCAATGAAGAAAGTACCTTACTG 1122
Db 361 AAGTATCACACATCCTTGATGAAGTACCGGATTCAGATTAAGGCTCCCAATGCGTGG 420
Qy 1123 GCTATGCGAATCTTGTGCAAGTCTATTTTCACTGTCGTTTAAACGCTATATAGTGTG 1182
Db 421 GCGTTTGAACACCGGCGAGAACCTTAATTTTACACTGCATGCACTGCAATGATGTAT 480
Qy 1183 CTCCGCACTGAGGAATTTGTGGCAGGTACCCCTTACTGTGTGAGATCTGATATGATG 1242
Db 481 ATGGCTTATATGCTGTATATGAGGCTCTCTTACAGTGGGGATCTTGTGTATTAAT 540
Qy 1243 GCACTGCTTTTACGCTTTCATTTACCCCTGAACTTTCTGGAACTGTATATAGAGACT 1302
Db 541 CAACCTGATTCACAGCTCTCGGTGCACTAACTTCTTGGTAGGCTTACCGTATCTC 600
Qy 1303 AGACAGACTCATGATATGATGAACCTTTGTACTTCTACTCAAGTGAACCAATTT 1362
Db 601 AAGAGTCTGTATGATATGGAATCTTTATTAACGTCAAAAATCAGGTCAATTT 660
Qy 1363 AAAAGCAAGATGATGCAATCTCCCTTGAGATCACACAGACAGCTACGGTGGCCCTT 1422
Db 661 AAGAACTCCCAAAATGCCAGAACCTTACCAATACAAACCGTGGATAT---TTGCTTT 717
Qy 1423 GATATGTCATTTTGAATATGATGAGGCGCAAGAACTCTTATGGAATATCTTTGAA 1482
Db 718 GAATATGTACGTTTGGCTATGACCCGAGGGGATATTTGAACAATGTTTCTTTACC 777
Qy 1483 GTCCCTGAGAGAAAGAGTGGCCATTTAGAGGATAGTGGCTCAGGAAAGCAATA 1542
Db 778 ATCCAGCTGGAATGAAACATGCAATGATAGGCCATCGGCTCGGGGAATCCACCAAT 837
Qy 1543 GTGAGGCTATATTTGCTTCTATGAGCCTCAAAAAGGTAGATTTATCTTGTGCTCA 1602
Db 838 TTGAAGCTCGATTTATGATTTCTATGAGCCGAGCAAGTCTGATTTCTTGTGGGCA 897
Qy 1603 AATATACAGATGTGAGCTGGAAGCCTTGCAGGGCAGTGGAGTGTACTCTAGAT 1662
Db 898 GATATCGCGATTTAGACTTCTTCTTACGAAAGCTATCGGTGTGCTCCCAAGAT 957
Qy 1663 GCTGCTCTTCTCATATATCTATTTATTAACAATCTTTATATGAAACATGATGCTTCA 1722
Db 958 ACTCCTCTTCTCAATGACACATCTGGAGATGTTAAATTCGGCAATATCAGTTCCCT 1017
Qy 1723 CCTGAGAGATGATGAGTGGCAAAATTTAGCTGCACTTCAATGATGCAATCTTCGATG 1782
Db 1018 GACGATGATTTCTCAGGCGCAATGAAAAAGCTCACTCAGAACTCATCAAGACCTA 1077
Qy 1783 CCACATGATATGACACCAAGTAGGGAGAGAGACTCAAGCTTCTGAGAGAAAG 1842
Db 1078 CCAAGGCGGCTTCCACCGTTGTAGGGAGCGCGTTGATGATCAGGGAGTGAGAA 1137
Qy 1843 CAAGAAGTGAATTCAGAGCAATTTTGAAGACCCCGCAATCATCTATGATGA 1902
Db 1138 CAAGGCTGTGATTTGCTGCTGTGCTTTTGAAGACGCTCGCTGATGTTTTCGAGAG 1197
Qy 1903 GCTACTTCATGCTTATGATGATTTACTGAGAGACTATCTTGTGCTCAT-----GAG 1956
Db 1198 GCTCAAGTGTCTTGAATACACACAGAGAGAGACTTTTGCACACATTCACAGAGAC 1257
Qy 1957 GATGTGTCAACACAGAACTTCTATTTTCAATGACACAGATTTGTCAACAGTGGTAT 2016
Db 1258 TTTTCTTCAATTCAGAGAGAGGTTTACGTTGGCCATATACCTGCGACAAATGCTGAT 1317
Qy 2017 GCAATGAATCATTTGCTTGTGATCAGGCTAAGGTAGCGCAACGTGTTACCCACATG 2076

Db 1318 GCAGATTAAGTATGTTGTTTGAACAAGTTCTGTCCGCGAAGGGCACACAGCTCG 1377
Qy 2077 TTGCTGCTAACCCCTCATATGATCTATTCAGAAATGTGGATACAGAGAGCC 2131
Db 1378 CTGTACGCTCAGAGAGATCCCTATACCGGGCTGTGTGGATATTCAGAAAAACC 1432

RESULT 4
US-09-134-001C-1893
; Sequence 1893, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1893
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1893

Query Match 8.1%; Score 188.8; DB 4; Length 1749;
Best Local Similarity 49.2%; Pred. No. 4.9e-45;
Matches 524; Conservative 0; Mismatches 537; Indels 3; Gaps 1;

Qy 1026 TGTGAATATTTAATATGAAGATATGAACACAGAGATATGATGATTTTGAAGAC 1085
Db 648 TATTAAGATTTGCTATGTGAAGCAATGAAGCTTAATAATTTGATTAACATTAACAGAA 707
Qy 1086 GTATGAGACTGCTTCAATGAAAAGTACCTTACTCTGTGCTATGCTGAATTTGTGCTAAG 1145
Db 708 TTTTTCACAGAGCTTCCACATACAGATGAGAGCATTTCTTTGTGCTATTTAA 767
Qy 1146 TGCTATTTTCAAGTGTGCTTTTAAAGCTATATATGCTGTGCTGCAAGTATGAGC 1205
Db 768 TACTGTATACAGATTTAGGCCCAATTAATGTGATGTGGGTGCTCATATTTTGGCAATTAC 827
Qy 1206 AGGTACCTTACTGTGAGATCTAGTATGATGATGATGATGATGATGATGATGATGATG 1265
Db 828 AGATGCTATTCGTGCGAACTCTAGCAGATTTGTGCTTATCTAGAACATATATTTGG 887
Qy 1266 ACCCTGAATCTTCTGGAAGCTGATATGAGAGACTAGACACTCATTATGATGA 1325
Db 888 ACCACTTGAAGACTAGTATCTTCACTTATCTACACTTACCAAGTTTGTGATATGGA 947
Qy 1326 CACCTGTTTACTCTTACTCAAGGTAGACACCAATTAAGAAAGCAATGATGCTATCC 1385
Db 948 CAGAGTATTCAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007
Qy 1386 CTTTCAATACACACAGACAGTACCGTGGCTTTGATATATGATGATTTGATATCAT 1445
Db 1008 AATTAATACATGAAGGTCAAAATGATTAACATGATGATGATGATGATGATGATG 1067
Qy 1446 TGAAGGCGAGAAAGTCTTATGATGATATCTTTGAAGCTCCCTGACAGAAAGAGTGG 1505
Db 1068 TGA---AAAAGATATTAACAGATATTAATTTAACAATTAACAAAGCGAACTGATAC 1124
Qy 1506 CATTTGAGAGTGTGCTCAGGAGAAAGACACATATGAGAGCTATATTTGCTTTCTA 1565
Db 1125 ATTTGATGATAGTGTGTGAGGAGAAATCTACTTGTATTAATCTTATACCAAGATTTTA 1184
Qy 1566 TGAAGCTCAAAAGGTAGATTTATCTGTGTGCAAAATATACAGATGAGAGCTGGA 1625
Db 1185 TGATGTACTCAAGGTGAATAATCTTATGATCATATTAATGATTAAGATTTCTTAATG 1244

Qy	1626	AAGCTTCGAGAGGGCAGTGGGAGTGGTAACTCAGAGATGTCCTCTTCATTAATCTAT	1685
Db	1245	TAGTTTAAGGAATCAATAGCCTTAGTACACAGATAAATATTTCTTTCTGATACGGT	1304
Qy	1686	TTATATCAACCTCTTATATGGAACATCAGTCTTCACTGAGAGAGTATGACATGGC	1745
Db	1305	TAAAGAGAAATTTTGTGTGGTAGGCTGATGACCATGATGATGAATCGTAGAAGCTGC	1364
Qy	1746	AAATTTAGCTGACCTTCATGATGCAATTTCTTCGAATGCGACATGAGATATATACGCCAAGT	1805
Db	1365	AAAAATGCGAATGCCATGATTTTATTTTCAATTTTACCGAATGGAATATATCTGAAGT	1424
Qy	1806	AGGGAGAGAGACTCAAGCTTTGAGAGAGAGAAACAAAGATAGCAATTCAGAGAGC	1865
Db	1425	AGGAGAACGAGAGATTAATATCTGCTGGACACAAAACAAAGTTGTCATTTGACAGTAT	1484
Qy	1866	CATTTTGAAGGACCCCGCAGTCATCTCTATGATGAAGCTTCTCATGTTAGATTGAT	1925
Db	1485	CTTTTAAATATATCCTCCTGTTTAAATTATGATGAAGCAACAAGTGCATTTGATTTAGA	1544
Qy	1926	TACTGAAGAGACTATCTTGATGCCATGAAGATGTGGTCAACACAGAACTTCTATTTT	1985
Db	1545	CGATBAACCTTATTTATTCAGAAGCACTTGATGTTTAAAGTADAGGATTAACCAACATTAAAT	1604
Qy	1986	CATTGCACACAGATTGTCAACAGTGGTTGATGCAGATGAATTCATTTGCTTGGATCAGGG	2045
Db	1605	TGTTTCACATCGTCTATCTTACCATTACTCAGCAGATAGGAATAGTTGTATGAGAAATGG	1664
Qy	2046	TAAAGTAGCCGACCGTGTATCCCAACCAATGTTTTCCTTGCTTAC	2089
Db	1665	ACGAATGTGTGAGCTGGCACACCAACCAATTAATTAATTAAC	1708

RESULT 5

US-08-961-527-63
Sequence 63, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P5340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 7760 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-961-527-63

```

Query Match      7.7%;   Score 180.2;   DB 4;   Length 7760;
Best Local Similarity 54.0%;   Pred. NO. 3.7e-42;
Matches 368;   Conservative 0;   Mismatches 313;   Indels 0;   Gaps 0;

```

OY	1394	TCACACAGACGACTAACCGCGCCTTGGATGAATGCGATTTGGATATACATATGAGGCC	1453
Db	6480	TCATCTAAGTTGCAGAAAGGTGTGAATATAGTCATATCGTTTTTCATACTGCTCGTATA	6539
OY	1454	AGAAAGCTCTTAGTGAATAATCCTTTGAAGTCCCTGCAGAAAGAAATGSCCATTTGAG	1513
Db	6540	AACCTATTTTGAAGAATGTCAGATTTTCGCCCCCTTAAGGCCAGATGACACAGATGTGG	6599
OY	1514	GAGGTAGTGGGTGAGGAAAAGCACAAATGATGAGGCTATTAATTTGCTTATAGACCTC	1573
Db	6600	GGCCGACAGGTTCAGGAAAAAGCATATATATGAACTCATCAATCGCTTTATGATGTGG	6659
OY	1574	AAAAAGGTAGCATTTTATCTTGCTGCTCAAAATATACAAAGATGTGAGCTGGAAGCCTTC	1633
Db	6660	ATGCTGTGTGATTTTATTTTGAATGGTAAAGACATTCGCGCATGACTCTTAATATGTCTTA	6719
OY	1634	GGAGGAGAGGGAGAGTGCCTACAGATGCTGCTCTCTCCATATATCTAATTAATACA	1693
Db	6720	GAAACAAAGGGGAATTGTTATTCAGATTCGATTCGTTGTTTAGCGGAACGATTAGAGACA	6779
OY	1694	ACCTCTATATGGAACATCAGTGCCTTCACTGAGAAAGTGTATGCACTGGCAAAATTTG	1753
Db	6780	ATATCGATTTGGTGTGGCCAGATGCTAAGTCAGAAATGTTGAGGTACACAGAAAAGCA	6839
OY	1754	CTGACATTCATGATGCAATTTCTTGAATGCCATGATATATACCCCAAGTAGGGGAAC	1813
Db	6840	CCCACTATTCACGATATATCGAAAGTTTCCCTGGAATAGACATCTCTTATATATGATG	6899
OY	1814	GAGGACTCAAGCTTTTACAGAGGAGAAACCAAGATAGTGCATTTGCAGAGACATTTGA	1873
Db	6900	ACCAAGACATCTTTTCAACGGGCAABACCAATGATTTCAATCGCTGAAACCTGATGA	6959
OY	1874	AGAGACCCCAAGTCATACTATATGATGAAAGCTAATCTTCAATCGTTAGATTGATCTGAG	1933
Db	6960	CAGATTCAGAAAGTTTCATCTTCGATGAAAGCAATTCMAACCTBATATACGTGACAGAAA	7019
OY	1934	AGACTATCTTGTGTCATGAGAAAGATGTGTCAAAACAGAACTTTCTATTTTCATTGAC	1993
Db	7020	GCAAGATTTCAAGCATGCCATGAGAGGTGGTGTGACAGGTATGATTTTCGTCATTTGCC	7079
OY	1994	ACGATTTGTCAACAGTGTGATGACGATGAATCATTTGTCTTGSATCAGGGTAAAGTATG	2053
Db	7080	ACCCCTTGAAAACCATTTCTCAATGACAGATCAATATATGTCTTAAAGATGAGAAATGCA	7139
OY	2054	CCGAACGTGGTACCCCAACCATG 2074	
Db	7140	TTTGACGTGTAAACCAACATG 7160	

RESULT 6

US-08-772-270A-6
Sequence 6, Application US/0872270A
Patent No. 6019984
GENERAL INFORMATION:
APPLICANT: MacInnes, Janet
APPLICANT: Ricciardi, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosenda, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada

ZIP: MSH 3V2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-6

Query Match 7.2%; Score 170; DB 3; Length 5120;
Best Local Similarity 55.1%; Pred. No. 2.8e-39;

Matches 332; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY 1502 TGGCCATTGTAGAGTGTGGGTGCGAGAAAGACAAATAGTGGCTATTATTTGCT 1561
DB 2084 TCGGATCGTAGAGCTTCAGGCTCAGGAGAGACCTTACGAAATATTTGACCTT 2143
QY 1562 TCTATGACCTCAAAAGGATGACATTATCTTCTGTCAAAATATACAGATGAGCC 1621
DB 2144 TTTTATATCCGAAACCGTACGATTAATGATGGCGCATGATTTAGCATTTGCCGATC 2203
QY 1622 TGGAAAGCCTTGAGGGGCGAGGTGACCTCAGAGATGCTGCTCTTCATATA 1681
DB 2204 CGAACTGCGTACGTCTTAAGTCGGGGGTGATTAACAAGATACGATTAATTAATCGTA 2263
QY 1682 CTATTTATTACACTCTTATATGGAACATCAGTGTCTTCACTGAGGAAGTATGACG 1741
DB 2264 GTATTCGAGATATATTCCTTAGCGGATCCGGGTATGCCAATGAAAAATTTGCCATG 2323
QY 1742 TGGCAAAATAGCTGACCTTATGATGCAATTTTGAATGCCAATGATATGACACC 1801
DB 2324 CGGCAAAATAGCCGCGCACATGAAATTTATTTCTGAATTCGTGAGGATATTAACGCA 2383
QY 1802 AAGTAAAGGGAACGAGACTCAAGCTTTCAGAGAGAAAAAGCAAGAGTATGCAAT 1861
DB 2384 TTGTGTGAGACAAAGTCCGGGGCTATCTGGCGGCAACGCAACGATTCGATGAC 2443
QY 1862 GAGCATTATTTGAGAGACCCCGCAGTCACTATGATGAAGTACTTATCATGTTAGATT 1921
DB 2444 GCGCTTTGTGTAATATCCGAAATCTTAATTTTGTGAAGCAACAGCGCATTTAGATT 2503
QY 1922 CGATTAAGAGACTATTTCTGTGTCATGAAGATGTGTCAAAACAGAACTTCTA 1981
DB 2504 ATGAATCCGAGCATATCATATGCGCAATATGACACAGATTTTAAAGGAGAAACGGTAA 2563
QY 1982 TTTTCATTGCAACAGATGTCACAGTGTGATGCAAGTAAATCAATTTGCTTGATC 2041
DB 2564 TTATCATTTGCAACCGCTTTATCTACGGTAAATAATGCGAACCGTATTAATGATGAAA 2623
QY 2042 AGGTAAGTATGCGAAGTGTACCAACCATGTTGCTTGTAAACCTCATAGTATCT 2101
DB 2624 AAGTCAAGATGTGGAACAGTAAAGCATTAAGAGCTGCTTGTGATCCAAACGGCTTAT 2683
QY 2102 AT 2103

DB 2684 AT 2685

RESULT 7

US-08-488-706-1
Sequence 1, Application US/08488706
Patent No. 5994525

GENERAL INFORMATION:
APPLICANT: Kamp, Elbarte M.

APPLICANT: Smits, Marius A.

TITLE OF INVENTION: Recombinant Vaccine For Prevention and/or

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: Abelman, Frayne & Schwab

STREET: 708 Third Avenue
CITY: New York

STATE: New York
COUNTRY: USA

ZIP: 10017
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC or compatible

OPERATING SYSTEM: DOS 3.31
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,706

FILING DATE: 09-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/722,971

FILING DATE: 28-JUNE-1991
ATTORNEY/AGENT INFORMATION:

NAME: Cinnamon, Jay S.
REGISTRATION NUMBER: 24,156

REFERENCE/DOCKET NUMBER: 201,875
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-949-9022
TELEFAX: 212-949-9190

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 8370
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

US-08-488-706-1

Query Match 7.2%; Score 170; DB 2; Length 8370;
Best Local Similarity 55.1%; Pred. No. 3.7e-39;

Matches 332; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY 1502 TGGCCATTGTAGAGTGTGGGTGCGGAAAGACAAATAGTGGCTATTATTTGCT 1561
DB 5334 TCGGATCGTAGAGCTTCAGGCTCAGGAGAGACCTTAAAGAAATATTAATCAACGTT 5393
QY 1562 TCTATGACCTCAAAAGGATGACATTATCTTGTGTCAAAATATACAGATGAGGCC 1621
DB 5394 TTTTATATCCGAAACCGTACGATTAATTAATGATGGCATGATTTAGCATTTGCCGATC 5453
QY 1622 TGGAAAGCCTTGAGGGGCGAGGTGATCCTGAGATGCTGTCTTCCATATA 1681
DB 5454 CGAACTGCGTACGTCTTAAGTCGGGGGTGATTAACAAGATACGATTAATTAATGTA 5513
QY 1682 CTATTTATTACACTCTTATATGGAACATCAGTGTCTTCACTGAGGAAGTATGACG 1741
DB 5514 GTATTCGAGATATATTCCTTAGCGGATCCGGGTATGCCAATGAAAAATTTGCCATG 5573
QY 1742 TGGCAAAATAGCTGACCTTATGATGCAATTTTGAATGCCAATGATATGACACC 1801
DB 5574 CGGCAAAATAGCCGCGCACATGAAATTTATTTCTGAATTCGTGAGGATATTAACGCA 5633
QY 1802 AAGTAAAGGGAACGAGACTCAAGCTTTCAGAGAGAAAAAGCAAGATGCAATTCGA 1861

Db 5634 TTGTTGTCAGCAAGTGCGGGGCTATCTGGGGGCAAGCCAGTATTTGGATTGCA.C. 5693
| | | | |
Qy 1862 GAGCATTGTTGAAGAACCCCGCATCTATGATGTAAGTACTTATGATTTAGATT 1921
| | | | |
Db 5694 GGGCTTTGGTGAATTAACCCGAAATCTTAAATTTTGTATGTAACGACCGAGCTTAGATT 5753
| | | | |
Qy 1922 CGATTACTGAAGAGACTATTTCTTGTGTCATGAGGATGTGTCAAACAGAACTTTCTTA 1981
| | | | |
Db 5754 ATGAATCCGAGATATCATCATGCGCAATATGCAACAGTTTGTAAAGGAGAAACGGTAA 5813
| | | | |
Qy 1982 TTTTCATTGACACAGATTGTCACAGTGTGATGACAGATGAATCATTTGCTTGGATC 2041
| | | | |
Db 5814 TTATCATTCACACCGCTTATCTACGGTAAATAATGCGCAGCTATTTATGATGAGAAA 5873
| | | | |
Qy 2042 AGGGTAAGTACCGCAACCTGTATCCCAACCATGTTTGTCTTAACCTCATAGTATCT 2101
| | | | |
Db 5874 AAGGTCAGATTGTGAACAAGGTAAGCATTAAGAGACTGCTTGTCTGATCCAAAGCGCTTAT 5933
| | | | |
Qy 2102 AT 2103
| |
Db 5934 AT 5935

RESULT 8

US-08-961-527-70
; Sequence 70, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13188 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-70

Query Match 7.1%; Score 165.8; DB 4; Length 13188;

Best Local Similarity 53.7%; Pred. No. 8,1e-38;

Matches 344; Conservative 0; Mismatches 297; Indels 0; Gaps 0;
Qy 1445 TTGAGGGCCAGAAAGCTTAGTGAATATCTTTGAAGTCCCTCGAGAAAGAAAGTGG 1504
| | | | |
Db 7421 TTGAAATGAGAAACACTGACGATATTCACCTTAGTTGGCAAAAGGCAAAACACTGG 7480
| | | | |

Qy 1505 CCATTGTAGAGAGTATGAGGTCAAGGAAAGACAAATAGTAGAGGCTTATTTGCTTCT 1564
| | | | |
Db 7481 GCTTGTTGGGCAACAGAGCTCTTGGAAGAAAGCTCTTAATCAAGCTCCTCTGTGTAAT 7540
| | | | |
Qy 1565 ATGAGCCTCAAAAGGGTAGCTTTATCTTGTGCTCAAAATATATCAAGATGTAGCCTGG 1624
| | | | |
Db 7541 ACGATGTGATTAAGGGTGCCTTATCTAAACGGTCAAGATATTCGGAGACTATGCTGTA 7600
| | | | |
Qy 1625 AAAGCTTCGGAGGGCAGTGGAGATGTGTAAGTCAAGATGCTGCTCTCCATTAATACTA 1684
| | | | |
Db 7601 CAGACTTCGACAGTCTATAGGCTATGTTCTCAAGACAGTTTCTTTTGGCACTTCAA 7660
| | | | |
Qy 1685 TTATTAACAACCTTTATATGAAGAAATCAAGTCTTCACTGAGAAAGTATGACAGTGG 1744
| | | | |
Db 7661 TCCAGACAAATATCGCTTGTGGCAATCCTTAACCTTGCCCTTTCAGCGGTGAGAGAGTA 7720
| | | | |
Qy 1745 CAAATTTAGCTGAGCTTCATGATGCAATTTCTGCAATGCCACATGATATGACACCCAG 1804
| | | | |
Db 7721 CTAAGCTAGCCCGGGTTTACCAAGATATTTGACATATGCTCAAGGATTTGATACGCTGA 7780
| | | | |
Qy 1805 TAGGGGAACGAGACTCAAGCTTTCAGAGAGAGAAAGCAAGAGTACATTCGACAGAG 1864
| | | | |
Db 7781 TTGGTGAAGAAAGAGATCAGTCTTTCTGTGTCAAAAGCAAGGTTGGCTATGAGTCGG 7840
| | | | |
Qy 1865 CCATTTGAAGGACCCCGCAGTCACTCTATGATGAGAGTACTTCAATCGTTAGATTGCA 1924
| | | | |
Db 7841 CTATGATTTTAGACCTCGTATATCTTGATTTTGATGATATCTTATCCGCGGTATATGCA 7900
| | | | |
Qy 1925 TTACTGAAGAGACTATTTCTTGTGTCATGAAGATGTGTCAACACAGAACTTTATTT 1984
| | | | |
Db 7901 AGACAGAGTATGCAATTAATGACAACCTCAAGGAGATGCCAAAGCAAGCAACCAATTA 7960
| | | | |
Qy 1985 TCATTGACACAGATGTGCAACAGTGTGATGATGAGATGAATATATGTCTGTGATACAG 2044
| | | | |
Db 7961 TCACGCCCATGCGCTCAGTGTGTGTTCATGACAGATTTATTTTGTCTACAAATG 8020
| | | | |
Qy 2045 GTAAGTAGCCGAAGGTGTACCAACCATGCTGTTGCTGTCT 2085
| | | | |
Db 8021 GTCAATTAATGAAACGAGGACGACGACGAAGACTTGCTAGCT 8061
| | | | |

RESULT 9

US-08-976-259-79/c
; Sequence 79, Application US/08976259
; Patent No. 6316609
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Choi, Gil H.
; APPLICANT: Welch, Rodney A.
; TITLE OF INVENTION: Nucleotide Sequence of *Bacterichia coli*
; Patent No. 6316609
; NUMBER OF SEQUENCES: 142
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,259
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.

```

;
;   REGISTRATION NUMBER: 36,688
;   REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
;   INFORMATION FOR SEQ ID NO: 79:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 3576 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;
US-08-976-259-79

Query Match      7.1%; Score 165.4; DB 4; Length 3576;
Best Local Similarity 53.7%; Pred. No. 5e-38;
Matches 340; Conservative 1; Mismatches 292; Indels 0; Gaps 0;

QY 1492 GGAAGAAAGTGGCCATTGTAGAGGTAGGCTCAGAGGAAACACAAATGTAGAGCTA 1551
DB 3564 GGGGAGGTTATGGTATGTCGACGTTCTGGTTCAAGAAAACACATTAATACTAAATTA 3505
QY 1552 TTATTTGCTCTTATGAGCCTCAAAAGGGTAGCATTTATCTTGGTGGTCAAAATATACAA 1611
DB 3504 ATTCAAGCTTTTATTTATCTCGAAAATGCGCAGCTTTAATTGATGACATGATCTTGGC 3445
QY 1612 GATGTGAGCCTGGAAGCCTTCGAGGGCAGTGGAGTGTACTCAGATGCTGTCTC 1671
DB 3444 TTGGCCGATCTTAAGTGTACGTGTCAGGTGGGGGTGTGTTGACAGCAATGTGCTG 3385
QY 1672 TTCCATAATCTATTTATTAACAACCTCTTATATGAAAACATCACTGCTTCACTGAGAA 1731
DB 3384 CTATATGCGACGTATTTATGTAAATATTTCACTGCTAATCTCGCATGCTCGTGA AAAA 3325
QY 1732 GTATATGAGTGGCAAAATTAAGTGAAGCTTATGATGCAATTTCTGAAATGCCACATGGA 1791
DB 3324 GTTATTTATGACGGAATTTATGACAGTGTCTATGATTTTATTTCTGAAATGCGTAGGGG 3265
QY 1792 TATGACACCCCAAGTAGGGAACGAGACTCAAGCTTTTCAAGAGAGAAAAGCAAGAGTA 1851
DB 3264 TATTAACCATTTGTGGGGAAACAGGGGGCAGAGATTATCCGAGGTCAACGTCAACGCATC 3205
QY 1852 GCATTTGCAAGACCATTTTGAAGACACCCCAAGTCACTCTTATGATGAAGCTATCTTCA 1911
DB 3204 GCATTTGCAAGGGCGTGGTGAACCAACCTTAATCTCACTTTGATGAAGCAACAGT 3145
QY 1912 TCGTATGATCGATTAATGAGAGACTATTTCTGTGCATGAAAGATGATGCTCAACAC 1971
DB 3144 GCTTGTGATTTATGATGAGAGCATGTCATCATGCGCAATATGACAAATATGTAAAGGC 3085
QY 1972 AGAATCTTCTATTTTCAATGACACAGATTTGCAACAGTGTGATGAGATGAATCATT 2031
DB 3084 AGAAGCTTATTAATCAATTTGTCATGCTGTCTACAGTAAATAATGACAGCCGATTAAT 3025
QY 2032 GTCTTGATCAGGGTAAAGTACCGAAACGTGTACCAACATGAGTTTCTTCTTCAACCT 2091
DB 3024 GTCATGAAAAAGGAAATTTGTGAACAGGGTAAACATTAAGAGCTGTTCTGAACCG 2965
QY 2092 CATGATCTATTGAGAAATGTGACATACACAG 2124
DB 2964 GAAAGTTTATACAGTTACTTATATACAGTTACAG 2932

RESULT 10
US-08-961-527-188
; Sequence 188, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
;   APPLICANT: Charles Kunsch
;   TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;   NUMBER OF SEQUENCES: 391
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
```

```

;
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,527
;
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 188:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 6492 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;
US-08-961-527-188

Query Match      6.6%; Score 154.8; DB 4; Length 6492;
Best Local Similarity 50.0%; Pred. No. 8.7e-35;
Matches 387; Conservative 0; Mismatches 387; Indels 0; Gaps 0;
```

```

QY 1301 CTAGACAGCACTCATATGATGAAACACCTGTTACTACTCAAGTAGACACCCAAA 1360
DB 2876 CTAAAGGGGCTTGATGAAATATTTATTTATGCGCCAAATAGAAAGATTAAAGAAA 2935
QY 1361 TTAAAGCAAGGATGAGCATCTCCCTTCAGATCAACACAGACAGCAACGCTGCGCT 1420
DB 2936 TTCAAAATCAAGATTTACAAAGGCAATGACTATGCTTAAATAATTTGATATGATC 2295
QY 1421 TTGATATGTCATTTTGAATATCATTTAGGGCCAGAAAGTCTTAGTGAATATCTTTG 1480
DB 2296 TAAAGATGTGAGTTTGCTTACATTAAGACCAAAAGTTTAAATGTGTAGTTTA 3055
QY 1481 AAGTCCCTGAGAAAGAAAGTGGCCATTGTAGAGATGTGGTCAAGGAAAGCAAA 1540
DB 3056 AAGCTAAGCAGGAGAGTCACTGCTTTGGTAGTGCAAGTGGCTCGGTAAACAACTA 3115
QY 1541 TAAGTGGCTATTTATTTGCTTATAGAGCTTCAAAAGGTAGACATTTATCTGTGCTG 1600
DB 3116 TCTTGAACCTTATATCAAGCTTTATGATTTATGACAAGGACAAATCTTAATGATGCA 3175
QY 1601 AAATATACAGATGTAGGCTGGAAGCCTTGCAGGGCAGTGGAGTGTACTCAGG 1660
DB 3176 AATATTAAGAAATATCAACGAATCCCTTTTGATTAAGGTGTCTATTTGTTCCAG 3235
QY 1661 ATGCTGTCTCTTCAATTAATCTATTTATTAACAACCTTTATATGAAACATCAGTGT 1720
DB 3236 ATGTGTTCTCTTAAATCAAAAGCTTATGAAATAATTAAGATCGGTAAAGCAAGTGA 3295
QY 1721 CACTTGAAGAGTGTATGACGTGGCAAAATTAAGTGTGATCTATGATGCAATCTTCGA 1780
DB 3296 GTGACGAGAGGTTAAAGGACGCAAACTTGCAATTTGACAGATTTTATGAAAAA 3355
QY 1781 TGCCATGAGATATGACACCAAGTAGGGGAAACGAGACTCAAGCTTTGACGAGAGAAA 1840
DB 3356 TGATATAAGTTTCAATACAGTATTTGTGAAACGAGAGTGAATATGAGAGAGAAA 3415
QY 1841 AGCAAGAGTAGCAATGCAAGAGCAATTTGAAAGACCCCGCATGCTATGATG 1900
```



```
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-461-823-1

Query Match      6.1%; Score 143.8; DB 1; Length 2726;
Best Local Similarity 52.9%; Pred. No. 8.6e-32;
Matches 335; Conservative 0; Mismatches 292; Indels 6; Gaps 1;

QY 1458 AGTCCTTAGTGAATATCTTTGAAGTCCCTGAGAAAGAAAGTGGCATTGTAGAGG 1517
    |||
DB 1457 AGTCTTAGGAGCTGAGCTCGAGGTGAAGAAGGCGCAGCGCTGCTGGGGAG 1716
    |||
QY 1518 TAGTGGGTGAGGAAAGCAATAGTGAAGCTATTATTTGCTTATGAGCCTCAAAA 1577
    |||
DB 1717 CAGTGGCTGTGGGAAAGACAGAGTGTCCAGCTCTGGAGCGGTTCTACGACCCCTTGGC 1776
    |||
QY 1578 GGGTAGCATTTATCTGTGCTGCAAAATATACAGATGTGAGCCTGGAAAGCCTTGGAG 1637
    |||
DB 1777 AGGAAAGTCTGCTGTGATGGCAAAATAAAGCAGCTGAATGTCTAGTCTCCAGC 1836
    |||
QY 1638 GGCAGTGGAGTGTGTACTCAGATGTGTCTCTCTCATATATCTATTATTAACAAC 1697
    |||
DB 1837 ACACCTGGGAGTGTGTGTCCAGAGGCCCATCTGTTGACGTGAGCATGTCTGAGAAC 1896
    |||
QY 1698 CTTATATGAAACATCAG-----TGCTTCACTGAGAGTGTATGCAATGSCAAAATT 1751
    |||
DB 1897 TGCCTATGAGACAAACAGCCGGTGTGTACAGAAAGATCTGTGAGGAGCAAAAGGA 1956
    |||
QY 1752 AGCTGACCTTCATGATGCAATTTCTTGAATGCCATGAGATATGACCCCAAGTAGGGA 1811
    |||
DB 1957 GGGCAACATATACCTCTTCATGAGTCACTGCTTAATTAATATAGCACTAAAGTAGAGA 2016
    |||
QY 1812 ACGAGACTCAAGCTTTCAGAGAGGAAAGCAAGATGCAATTCAGAGGCCATTCTT 1871
    |||
DB 2017 CAAGGAACCTAGCTCTCTGTGTGCGCAGAAACAGCATTCCTAGCTCTGTGCTTGT 2076
    |||
QY 1872 GAAGAGCCCCCACTCATCTATGATGAAGTACTTTCATGCTTGAATTCGATTACTGA 1931
    |||
DB 2077 TAGACAGCTCATATTTTGTGTTGATGAAGCCAGCTCTGATACAGAAAGTGA 2136
    |||
QY 1932 AGAAGACTATTTGTGTGTCATGAAGATGTGTCAAAACAGAACTTCTATTTCATTGC 1991
    |||
DB 2137 AAAGGTGTCCAAAGAGCCCTGGAACAAGCCAGAGGCCGACCTGTGCAATTTGATTGC 2196
    |||
QY 1992 ACACAGATTGTCAACAGTGTGTGATGAGATGAATCATTTGCTTGGATCAGGTTAAGT 2051
    |||
DB 2197 TCACCGCTGTCCACATCCAGAAATGCAACTTAATAGTGTGTTTCAGAAATGCAAGT 2256
    |||
QY 2052 AGCCGAACGTGTACCCACCATGTGTTGCTTGC 2084
    |||
DB 2257 CAAGAGCATGTGACGATCAGACAGCTGCTGC 2289
    |||

RESULT 13
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
```

```
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg. No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-1

Query Match      6.1%; Score 143.8; DB 2; Length 4264;
Best Local Similarity 52.9%; Pred. No. 1.1e-31;
Matches 335; Conservative 0; Mismatches 292; Indels 6; Gaps 1;

QY 1458 AGTCCTTAGGATATATCTTTGAAGTCCCTGAGAAAGAAAGTGGCATTGTAGAGG 1517
    |||
DB 3291 AGTCTTAGGAGCTGAGCTCGAGGTGAAGAAGGCCAGCCCTGCTGTGGGGAG 3350
    |||
QY 1518 TAGTGGGTGAGGAAAGCAATAGTGAAGCTATTATTTGCTTATGAGCCTCAAAA 1577
    |||
DB 3351 CAGTGGCTGTGGGAAAGACAGTGTGTCCAGCTCTGAGCGGTTCTACGACCCCTTGGC 3410
    |||
QY 1578 GGGTAGATTATCTTGTGTGTGATGAAGTATGAGATGTGAGAGCCTGGAAGCCTTGGAG 1637
    |||
DB 3411 AGGAAAGTGTCTCTGATGAGCAAAAGAAATGAAGCACTGAATGTTCACTGCTCGAGC 3470
    |||
QY 1638 GGCAGTGGAGTGTGTACTCAGATGTGTCTCTTCATATATCTATTATTAACAAC 1697
    |||
DB 3471 ACACCTGGGAGTGTGTGTCCAGAGGCCCATCTGTTGACGTGAGCATTTCTGAGAAC 3530
    |||
QY 1698 CTTATATGAAACATCAG-----TGCTTCACTGAGAGTGTATGCAAGTGCAGAAATT 1751
    |||
DB 3531 TGCCTATGAGAGCAACAGCCGGTGTGTGTCAGAAAGATTTGTAGGCGCAAAAGGA 3590
    |||
QY 1752 AGCTGACCTTCATGATGCAATTTCTTGAATGCCATGTGATATGACCCCAAGTAGGGA 1811
    |||
DB 3591 GGGCAACATATACATGCTTCACTGAGTCACTGCTTAATTAATATAGCACTAAATAGAGA 3650
    |||
QY 1812 ACGAGACTCAAGCTTTCAGAGAGGAAAGCAAGATGCAATTCAGAAAGCCATTCTT 1871
    |||
DB 3651 CAAGGAACCTAGCTCTCTGTGTGCGCAGAAACAGCATTCCTAGCTGTGCTTGT 3710
    |||
QY 1872 GAAGAGCCCCCACTCATCTATGATGAAGTGTGTCAAAACAGAACTTCTATTTCATTGC 1931
    |||
DB 3711 TACACAGCTCATATTTTGTGTTGATGAAGCCAGTCACTGCTGATACAGAAAGTGA 3770
    |||
QY 1932 AGGACTATTTCTGTGTCATGAAGATGTGTCAAAACAGAACTTCTATTTCATTGC 1991
    |||
DB 3771 AAAGGTGTCCAAAGAGCCCTGGAACAAGCCAGAGAGGCCGACCTGTGATTTGC 3830
    |||
QY 1992 ACACAGATTGTCAACAGTGTGTGATGAGATGAATCATTTGCTTGGATCAGGTTAAGT 2051
    |||
DB 3831 TCACCGCTGTCCACATCCAGAAATGCAACTTAATAGTGTGTTTCAGAAATGCAAGT 3890
    |||
QY 2052 AGCCGAACGTGTACCCACCATGTGTTGCTTGC 2084
    |||
```


Db 3891 CAAGAGCATGCGACGATCAGCAGCTGCTGCG 3923

RESULT 14

US-08-784-649A-5
Sequence 5, Application US/08784649A
Patent No. 5830697

GENERAL INFORMATION:

APPLICANT: Sikic, Branimir I
APPLICANT: Chen, Gang
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
CYCLOSPORIN MODULATION
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-784-649A-5

Query Match 6.1%; Score 143.8; DB 2; Length 4264;
Best Local Similarity 52.9%; Pred. No. 1.1e-31;
Matches 335; Conservative 0; Mismatches 292; Indels 6; Gaps 1;

Oy 1458 AGTCCTTAGTGAATATCCTTTGAGTCCCTGCAGGAAGAAAGTGGCATTGTAGGAGG 1517
Db 3291 AGTGTCTCAGGAGCTGAGCTGAGAGTGAAGAGGCGCAGCCTGCTGTGGGAG 3350
Oy 1518 TAGTGGGTGAGGAAGAAAGCAATAGTGAAGCTATTATTCGCTTCTAGAGCTCAAAA 1577
Db 3351 CAGTGGCTGTGGGAAGAGACAGTGTCCAGCTCTGAGCGGTCTAGAGACCTTGGC 3410
Oy 1578 GGGTGCATTTATCTTGTGTCGCAAAATATACAGATGTAGCTGAGAAAGCTTCGGAG 1637
Db 3411 AGGGAAGGTCTGTGATGAGCAAAATAAAGCAAGTATGTTCAAGTGGCTCCGAGC 3470
Oy 1638 GGCAGTGGAGGTGACCTCAGAGATGCTCTTCCATAATATTAATAAAGT 1697
Db 3471 ACACCTGGGCACTGTGCTCCAGAGCCCATCTGTTGATGACACATTGTCGAGAACT 3530
Oy 1698 CTTATATGAAACATCAG-----TGCTTCACTGAGAGAGTATGACAGTGGCAAAATT 1751
Db 3531 TGCCTATGAGACCAACAGCGGGGTGTGTCAAGAGAGATTTGAGGGCGAGCAAGGA 3590
Oy 1752 AGCTGGACTTCATGATGCAATTTCTTGAAATGCAACATGAGATGACACCAAGTAGGGA 1811
Db 3591 GGCCTAACATACATGCTTCATCAGATGCTGCTTAATAATATACACTAAAGTAGGAGA 3650

Oy 1812 ACAGAGACTCAAGCTTTACAGAGAGAGAAAGCAAGATGCAATTCAGAGCCATT 1871
Db 3651 CAAGGAACATGAGCTCTGTGTGGCCAGAAACAGCATTCAGTCTGCTGCTGT 3710
Oy 1872 GAAGAGACCCCACTGATCTATGATGAAGTACTTATGTTAGATTGATTAAGT 1931
Db 3711 TAGACAGCTCATATTTGCTTTGGATGAAGCCACGCTGATGATACAGAAAGTGA 3770
Oy 1932 AGACATATTTGTCGATGCAAGAGATGTGTGCAACACAGAACTCTATTTTCATTGC 1991
Db 3771 AAAGTGTCCAGAGAGCTTGGACAAAGCCAGAGAGCCGACCTGCAATGTGATTGC 3830
Oy 1992 ACACAGATTGTCAACAGTGTGTTGATGAGATGAATCATTTGTTGATCAGGTTAAGT 2051
Db 3831 TCACCGCTGTCCACATCCAGATGAGACTTAATAGTGTGTTTCAGATGCGAGAGT 3890
Oy 2052 AGCCGAACGTGTACCCACCATGCTTTGCTTGC 2084
Db 3891 CAAGAGCATGCGACGATCAGCAGCTGCTGCG 3923

RESULT 15

US-08-181-471-2
Sequence 2, Application US/08181471
Patent No. 5641508

GENERAL INFORMATION:

APPLICANT: Li, Lingna
APPLICANT: Lienko, Valeryi K.
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
COMPOSITIONS TO HAIR FOLLICLES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: Thomas Fitting
STREET: 12526 High Bluff Drive, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92130

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/181,471
FILING DATE: 13-JAN-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/041,553
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: ANT0029P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 4646 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS
LOCATION: 425..4267

US-08-181-471-2

Query Match 6.1%; Score 143.8; DB 1; Length 4646;

Beat	Local Similarity	52.9%	Pred. No. 1.2e-31,
Matches	335;	Conservative	0; Mismatches 292; Indels 6; Gaps 1;
Qy	1458	AGTCCTTAGTGAATAATCCTTTGAAGTCCCTCAGAGAAAGAAAGTGCCATTGTAGAGG	1517
Db	3577	AATGCTTAGGGAACTGAGCTGAGGTGTAAGAAAGGCCAGACGCTGCTCTGTGGCAG	3636
Qy	1518	TATGGGTCAGGAAAGACAATAAGTAGGCTATATTGCTTCTATAGACCTCAAA	1577
Db	3637	CAGTGCGCTGGGAAAGACAGAGTGCTCAGCTCCGAGCGGTTCTAGACCCCTTGGC	3696
Qy	1578	GGGTAGCATTTATCTTGCGGTCAAAATATACAGATGAGCCCTGGAAAGCCTTGGAG	1637
Db	3697	AGGAAAGTGCTGCTTGTATGGCAAGAAATTAAGGCACTGAATGTTCACTGCTCCGAGC	3756
Qy	1638	GGCAGTGGAGTGTACCTCAGAGATGCTGCTCTCTTCATAATACTATTATTAACAACCT	1697
Db	3757	ACACCTGGGCATCGTGTCCAGAGGCCATCTGTGTTGATCTCAGATTTGCTGAAACAT	3816
Qy	1698	CTTATATGAAACATCAG-----TGCTTACCTGAGGAATGTATGCACTGGCAAAATT	1751
Db	3817	TGCCATATGAGAACAAAGCGGGGTGTGTCAAGAAAGATCGTAGGGCCAAAGGA	3876
Qy	1752	AGCTGAGCTTCATGATGCAATTCTTGAAATGCAATGSAATATGACACCCAGTAGGGGA	1811
Db	3877	GGCCAACTATACATGCTTCATCTGACTGACTGCTATTAATTAATAGCACTAAAGTGGAGA	3936
Qy	1812	ACGAGACTCAAGCTTTCAGAGAGAAAGCAAAAGATGAGCAATTCGACAAGCCATTTT	1871
Db	3937	CAAAAGAACTACGCTCTCTGTGGGCCAGAAACAAAGCATGCGATAGCTCGTGGCCCTTGT	3996
Qy	1872	GAGGAGCCCCAGTCATACTCTATGATGAGACTACTTCATCTGTTAGATTGATTACTGA	1931
Db	3997	TAGACAGCCTCATATTTTGGCTTTTGGATGAAACCACGTCAGCTCTGATATACAAAGTGA	4056
Qy	1932	AGAGACTATTTCTGGGCGCATGAAGATGTGTCAAACACAACTTCTATTTTCAATGC	1991
Db	4057	AAAGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAAAGCCGACCTGCAATTGTATTGC	4116
Qy	1992	ACACAGATTGTCACAGTGTGATGTCAGATGAAATCAATTGCTTGGATTCAGAGGTAAAGT	2051
Db	4117	TCACGCGCTGTCCACATCCATCCAGAAATGACAACTTAATAGTGTGTTTCAGAAATGGACAGT	4176
Qy	2052	AGCCGAAGTGTATCCCACTCATGTTTGGTTGC	2084
Db	4177	CAAGAGATGCGACGCATTCACAGACTGCTGGCC	4209

Search completed: February 15, 2003, 03:23:21
Job time : 164 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 03:21:35 ; Search time 121 Seconds
(without alignments)
9870.777 Million cell updates/sec

Title: AF133659
Perfect score: 2345
Sequence: 1 ATGCGCTGCTCGCATGCA.....TATTAATAATCATCATT 2345

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	358.6	15.3	5369	9 US-10-076-157-3	Sequence 3, Appli
2	282.6	12.1	1933	10 US-09-953-688A-4	Sequence 4, Appli
3	185.8	7.9	939	10 US-09-974-300-6511	Sequence 6511, Ap
4	184.6	7.9	1737	10 US-09-815-242-8403	Sequence 8403, Ap
5	182.4	7.8	1716	10 US-09-815-242-6522	Sequence 6522, Ap
6	165.4	7.1	3576	10 US-09-956-004-79	Sequence 79, Appli
7	158.2	6.7	1233	10 US-09-974-300-2176	Sequence 2176, Ap
8	157.8	6.7	6415	10 US-09-070-927A-275	Sequence 275, App
9	157.4	6.7	3924	10 US-09-880-107-2299	Sequence 2299, Ap
10	155	6.6	3601	10 US-09-070-927A-255	Sequence 255, App
11	154.6	6.6	1791	10 US-09-974-300-2172	Sequence 2172, Ap
12	154	6.6	9047	10 US-09-070-927A-12	Sequence 12, Appli
13	153.2	6.5	3861	9 US-09-938-842A-2237	Sequence 2237, App
14	152	6.5	3861	9 US-09-938-842A-263	Sequence 263, App
15	151.8	6.5	1863	9 US-09-917-800A-1560	Sequence 1560, Ap
16	151.8	6.5	2320	10 US-09-953-259-1	Sequence 1, Appli
17	149.4	6.4	4480	10 US-09-070-927A-320	Sequence 320, App
18	149.2	6.4	2681	10 US-09-749-340-8	Sequence 8, Appli
19	147.6	6.3	10391	10 US-09-070-927A-133	Sequence 133, App

20	147.4	6.3	4317	9 US-10-044-671-1	Sequence 1, Appli
21	147.2	6.3	1966	10 US-09-070-927A-433	Sequence 433, App
22	146.4	6.2	5010	10 US-09-917-800A-483	Sequence 483, App
23	146	6.2	3512	10 US-09-749-340-7	Sequence 7, Appli
24	146	6.2	3912	10 US-09-917-800A-1560	Sequence 1560, Ap
25	145.2	6.2	2066	10 US-09-873-409-9	Sequence 9, Appli
26	145.2	6.2	2856	10 US-09-873-409-10	Sequence 10, Appli
27	145.2	6.2	3177	10 US-09-873-409-12	Sequence 12, Appli
28	145.2	6.2	3621	10 US-09-873-409-14	Sequence 14, Appli
29	145.2	6.2	3702	10 US-09-873-409-13	Sequence 13, Appli
30	144	6.1	1761	9 US-10-260-877-25	Sequence 25, Appli
31	144	6.1	1754	10 US-09-815-242-6863	Sequence 6863, Ap
32	144	6.1	4175	10 US-09-749-340-3	Sequence 3, Appli
33	143.8	6.1	3860	10 US-09-866-866A-1	Sequence 1, Appli
34	143.8	6.1	3860	10 US-09-866-866A-3	Sequence 3, Appli
35	143.8	6.1	4643	9 US-10-072-621-2	Sequence 2, Appli
36	143.8	6.1	8630	10 US-09-306-417-2	Sequence 2, Appli
37	143.6	6.1	2698	10 US-09-749-340-5	Sequence 5, Appli
38	142.6	6.1	12438	10 US-09-070-927A-173	Sequence 173, App
39	142.6	6.1	1800	10 US-09-974-300-6492	Sequence 6492, App
40	142.4	6.1	7296	10 US-09-070-927A-59	Sequence 59, Appli
41	142.2	6.1	8630	10 US-09-306-417-1	Sequence 1, Appli
42	139.4	5.9	1201	10 US-09-765-272-189	Sequence 189, App
43	137.4	5.9	4189	10 US-09-866-866A-5	Sequence 5, Appli
44	135.8	5.8	981	10 US-09-770-445-272	Sequence 272, Appli
45	135.4	5.8	1194	10 US-09-974-300-2187	Sequence 2187, Ap

ALIGNMENTS

RESULT 1
US-10-076-157-3
; Sequence 3, Application US/10076157
; Publication No. US20030027309A1
GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Suelberger, Harald
; APPLICANT: Joefferen, Hans Wolfgang
; APPLICANT: Doval, Jose Luis Revuelta
; APPLICANT: Jimenez, Alberto
; APPLICANT: Garcia, Maria Angeles Santos
; TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii and the use th
; TITLE OF INVENTION: 1n
; FILE REFERENCE: 48684D1V
; FILE REFERENCE: microbial riboflavin synthesis
CURRENT APPLICATION NUMBER: US/10/076,157
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 09/212,247
PRIOR FILING DATE: 1998-12-16
NUMBER OF SEQ ID NOS: 21
SOFTWARE: WordPerfect v. 6.1
SEQ ID NO 3
LENGTH: 5369
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..1482
; NAME/KEY: CDS
; LOCATION: 1767..3299
; NAME/KEY: CDS
; LOCATION: 3588..4703
US-10-076-157-3

Query Match 15.3%; Score 358.6; DB 9; Length 5369;
Best Local Similarity 54.1%; Pred. No. 9e-90;
Matches 777; Conservative 0; Mismatches 649; Indels 9; Gaps 2;

CY 703 AACCTGATGCTTTTCCTGACGACACACCGGACCTTATCTAGGCTATTGAC 762
DB 1 AACCTGATGCTTTTCCTGACGACACACCGGACCTTATCTAGGCTATTGAC 60

```

QY 763 AGAGAAACAAGGGATATCAGTTTGTCTGAGTCTTTGGTATTAATCTTCCATC 822
Db 61 CTGGTTGAAAGTATCTTATGTGCTCAGTGCATGGTTTTTTCATATACGAT 120
QY 823 ATGTTGAAGTATGCTTGTCTGAGTGTGTTTGTATTAATAATGGCGGCCAGTTGT 882
Db 121 AATTTGAAATATCAATGGATATGTCATATGACATACAGTTGGTCTTCTCGCT 180
QY 883 TTGTAACCTTGGACACATTTGTCATATCAGACATTCACAGTTGCATCAGCGT 942
Db 181 GCTATTAACCTTTCGATATGCTTCTTATCTCATCTTACTTTCAGAACGAGCGT 240
QY 943 AGACTATGATTTAGAAATAGAAATGAAACAAGATATGATGACAGTATGCTGTATA 1002
Db 241 GCAACAGGTTTAAAGGCTGATGCAAGCAAGCTGACAAATTAAGCGCTAGTGGCAT 300
QY 1003 GACTACGCTGATATTAATAAATCTGAAAGTATTTTAATAATGAAAGATATGACAG 1062
Db 301 GATTCCTTAATTAATTTTGAAGCTGTAAAGTATTTCAATACGAGAGTACCTTGGGAC 360
QY 1063 AGATATGATGATTTTGAAGATATGAGACTGCTTCAATGAAAAAGTACTCTACTG 1122
Db 361 AGATATCAACATCTTATGATGAAATGACGGGATTCACATTAAGTCTCGCAATGCTG 420
QY 1123 GCTATGCTGAATTTGTGTCAAAGTCTATTTTCAGTGCGTTTACAGCTTAATGCTG 1182
Db 421 GCGTTTGAACACCGGCGCAAGACCTAATTTTACACATGCACTGCTCAATGATAT 480
QY 1183 CTGCGCAGTCAGGSAATTTGGCAGGTACCTTACTGTGGAGATCTAGTAATGTTGAT 1242
Db 481 ATGGCTGTAAATGGTGTATGAGAGGCTCTTCAACGTGGGATCTTGTATTAATAT 540
QY 1243 GGACTGCTTTTCAAGCTTCAATTAACCTTGAATCTTCTGGAACTGTATATAGAGACT 1302
Db 541 CAATCTGATATTCAGAGTCTCCGTCGACATAAATCTTCTGTGATGCTTCACTGATCTC 600
QY 1303 AGACAGACCTATATATATGAAACACTTGTATCTTACTAATCAAGGTAGACACCAAT 1362
Db 601 AAGCAGTCTCTATATGATATGAAATCTTATTTAACTCAAAAAAATCAGTCACAAT 660
QY 1363 AAAGCAAAAGTATGCACTCTCCCTTCAGATCACCAACACAGCTACCGTGCTT 1422
Db 661 AAGAACTCCCAATGCCCCAGAACCTACCAATACCAAAACCGTGTGATA--TTGCTTT 717
QY 1423 GATTAATGTCATTTTGAATATACATTTGAGGSCCAGAAAAGTCTTATGTAATCTTTGAA 1482
Db 718 GAAATGTATGCTTTGGCTATGACCCGAGCGCGCTATATTAACAATGTTCGTTTACC 777
QY 1483 GTCCCTGCAGAAAGAAAGTGGCCATTTAGAGAGTATGGGTCAAGGAAAAAGCAAT 1542
Db 778 ATCCCAAGCTGAAATGAAAGCTGCATATGAGCCCATCGGGCTCGGGGAAGTCCACAT 837
QY 1543 GTGAGGCTATATTTGGCTTCTATGAGCTCAAAAGGATGCAATTTATTTGCTGTCA 1602
Db 838 TTGAAGCTGTATATGATTTCTATGAGCCGAGCAAGGTGATCTTACTTGGCGGACA 897
QY 1603 AATATCAAGATGTAGGCTGGAAGCCTTGGAGGGGAGTGGAGTGTACTACAGAT 1662
Db 898 GATATCCGCGATTTAGACTTGTCTTCTTATACGAAAGCTATGCGTGTGCGGCCAAGAT 957
QY 1663 GCTGCTCTTCCATTAATCTATTTATTAACAACCTTATATGAAACATCAGTGTCA 1722
Db 958 ACTGCTCTTCAATGACATCTGGAGAAATGTTAAATTTGGGCAATATCTGTTCTCT 1017
QY 1723 CTTGAGGAAGTATGACAGTGGCAAAATTAAGTGAATTCATGATGCAATTTCTGCAAT 1782
Db 1018 GACGATGAAATTTCTCAGGCGCATAGAAAAAGCTCAACTCAGAACTACTCCAGAACT 1077
QY 1783 CCACATGATATGACACCCAGTAGGGGAAACGAGACTCAAGCTTTCAGAGAGAAAG 1842
Db 1078 CCAAAAGGCGCTTCAACCGTGTAGGGAGCGCGGTTTATATATCAGCGGAGTGAATA 1137
QY 1843 CAAGAATGACAAATTTGCAAGAGCCATTTTGAAGACCCCCAGTCAATCTTATGATGA 1902

```

```

Db 1138 CAAGGCTTGTATTTGCTCGTGTCTTTTGAAGACGCTCCGCTGATTTTTGACAG 1197
QY 1903 GCTACTCATGTTAGATTCATCTGAAAGACTATTTCTGGTCCAT-----GAAG 1956
Db 1198 GCTACAAAGTCTCTGATATACACACAGAGAGGACCTTTCACACCATTCACAGAAC 1257
QY 1957 GATGTGTCAACACAGAACTTTATTTTCAATTCAGACAGATTTGTCAAGTGTGAT 2016
Db 1258 TTTTCTTCAATTCAAAGACAGCGTTTACGTTGCCATATGACTGCGCAATCGCTGAT 1317
QY 2017 GCAGATGAATCATTTGTTGATTCAGGTTAAGTAAAGTAAAGTGTATCCCATGAT 2076
Db 1318 GCAGATGAATCATTTGTTGATTCAGGTTAAGTAAAGTGTAAAGTGTAAAGTGTAA 1377
QY 2077 TTGCTGTAAACCTTCATATGATATCTATTCAGAAATGTGCAATACACAGAGAGCC 2131
Db 1378 CTGTAGGCTACAAAGATCCCTATATCCGGGCTGTGTGGATATTCAGAAAAACC 1432

```

RESULT 2

US-09-953-688A-4

Sequence 4, Application US/09953688A

Patent No. US20020102649A1

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Yee, Henry

APPLICANT: Reddy, Roopa

APPLICANT: Gorgone, Gina

APPLICANT: Corley, Neil C.

APPLICANT: Azimzal, Yalda

APPLICANT: Patterson, Chandra

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: HUMAN TRANSPORT PROTEIN HOMOLOGS

FILE REFERENCE: PP-0555 US

CURRENT APPLICATION NUMBER: US/09/953,688A

PRIOR FILING DATE: 2001-09-12

PRIORITY FILING DATE: 1998-07-10

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 1933

TYPE: DNA

ORGANISM: HOMO SAPIEN

FEATURES:

OTHER INFORMATION: 2074412, ISITNOT01

US-09-953-688A-4

Query Match 12.1%; Score 282.6; DB 10; Length 1933;

Best Local Similarity 50.5%; Pred. No. 9.1e-69;

Matches 743; Conservative 0; Mismatches 79; Indels 9; Gaps 2;

```

QY 666 TCTTTCTCATCTTCAACAACCTGATCTGGGTTTTCACCTGACAGACAGCGGAGCTT 745
Db 269 TCTTCTCCACCTGACAGCTTCACTGCGCTGGCACTTGGGCGCGCACAGGGAGG 328
QY 746 TATCTAAGGCTATGACAGAGAAACAAGGGTATCAAGTTTGTCTGAGTCTTGTAT 805
Db 329 TGTGCGGATGCGGATGCGGACATCAAGTGTACAGGCTGTCTGACTACCTGAT 388
QY 806 TTAATCTTCTCCATCATCTGTTTGAAGTATGTCAGTGTGTTTGTATTAACAAT 865
Db 389 TCAATGTATCCCAACGCTGGCCGACATCATCTTGGCATATCTTACAGCATGTTCT 448
QY 866 GCGGTGCCAGTTGCTTGTGTAACCTTGAACAATTTGTATCATACACAGATTCACAG 925
Db 449 TCAAGCTGTGTTGGCTCTATGTTCTGTGATAGTCTTACCTTCAACCTGACCA 508
QY 926 TTGCACTACACGCTGGAAGTATGATTTAGATTAAGAAATGAACAAGCAGATATGATG 985
Db 509 TTGTGTCACTGAGTGGAGAACAAAGTTTGTCTGTCTATGAACACACAGAGAAAGCTTA 568

```

QY 986 CAGTAAATGCTGCTAATACTACCTGCTGAATTAATGAACCTGTGAATTTTAATATG 1045
Db 569 CCGGGGCAAGAGCTGAGTCTCTCTCTAACTTCAGACGGTGAAGATTAAACAACGCCG 628
QY 1046 AAAGATATGAACACAGAGATATGATGATTTTGAAGACGATAGAGCTGCTTATGA 1105
Db 629 AGAGTTAGAAAGTGAACGCTATCGAAGGCCATCAAAATATCAGGGTTTGAAGTGA 688
QY 1106 AAAGACTTACTCTGCTGCTATGCTGAACCTTGTCAAAGTCTAATTTTCACTGCTG 1165
Db 689 AGTGAGGCTTCACTGCTTTTACTAATTCAGACCCAAACCTGCTGATTTGGCTCGGAC 748
QY 1166 TAAACGCTAATATGCTGCTCCGCACTCAGGGAATGTGACAGTACCTTACTGTGAG 1225
Db 749 TCTCGCGGCTCCCTGCTTGGCGATCTTGTCTACAGACAGAAAGCTACAGGTGGG 808
QY 1226 ATCTAGTAATGTAATGAGCTGCTTTTCACTTCACTTCCCTGAACTTTCTGGAA 1285
Db 809 ACTATGCTCTTTGGGCACTTACATTAATCAGCTGTACATGCCCTCAATGTGTTGCA 868
QY 1286 CTGTATATAGAGACTAGACAGCACTCATATGATATGAAACACTTGTTTACTCTCA 1345
Db 869 CCTACTACAGATATTCAGACCACTTCACTGACATGGAACATGTTTACTTCTGCA 928
QY 1346 AGGTAGACCCCAATTAAAGACAAAGTATGCACTCCCTTCAATCAACACACAGA 1405
Db 929 AAGGAGACAGAAAGTAAAGACCTTCTGAGACAGGCCCCCTTGGCTT-----TCGA 982
QY 1406 CAGTACCGGCTGCTGATATGCACTTTGAATATATGAGGCGCAAGAACTCTTA 1465
Db 983 AGGCGCTATGAGTTGAGAACGCGCACTTACGTAAGCGAGGGGAGACTGTC 1042
QY 1466 GTGAAATATCTTTGAATGCTCCCTGACAGAAAGAAAGTGGCAATGTAGAGTATG 1525
Db 1043 AGAGGTCTCTTTCACTGTATGCTGACAGACACTTGGCTGTGGCCCACTTGGGG 1102
QY 1526 CAGGAAAGACAAATATGAGGCTATTTTCCGCTTATAGCCTCAAAAGGTAGCA 1585
Db 1103 CAGGAAAGACCAATTTTGGCGCTGCTTCCCTTCAAGATCACTCTGGCTGCA 1162
QY 1586 TTTATCTTGTGCTGCTCAAAATATACAGATGTAGCTGGAAGCCTTGGAGGGCACTG 1645
Db 1163 TCCGAATAGATGGGACAGACATTTCAAGGAGCCCAAGGCTCTCCGCTGTCACATG 1222
QY 1646 GAGTGTACTCTCAGAGTGTCTCTCTCAATTAATTAATTAACCTCTTATATG 1705
Db 1223 GAGTGTGCCCCAAGACACTGCTCTTTATATGACCATGCCGACATATCCGTTACG 1282
QY 1706 GAAACATCAGTCTTCACTGAGGAAGTATGCAATGGCAAAATTAAGTGAATTCATG 1765
Db 1283 GCCGTGTCAAGCTGGGAATGATGAGGTGAGGCTGCTGCTCAGGCTGCAAGGATTCATG 1342
QY 1766 ATGGAATCTTCCGAATGCCATGTGATATGACACCAAGTATGGGAAACGAGATCTAAGC 1825
Db 1343 ATGCAATTATGCTTCTTCCGAAAGGTACAGAGACAGGTGGGACCGGGACTTGAAGC 1402
QY 1826 TTTGAGAGGAGAAACCAAGATAGCAATTTGCAAGCCATTTTGAAGACCCCCAG 1885
Db 1403 TGAAGCGCGGGAGAACAGCGGCTCCGCACTTCCGACACATCTCAAGGCTCGGGA 1462
QY 1886 TCATATCTATGATGAAGTACTTCACTGTTAGTGTGATTAATTAAGAGACTATTTG 1945
Db 1463 TCATTTCTGTGATGAGGCAACGTCAAGGCTGATTAATTAATGAAGGGCCATTCAGG 1522
QY 1946 GTGCAATGAAGATGTGTCAAACAGAACTTATTTTCACTGCAACAGATTGTCAA 2005
Db 1523 CTCTTCCGCAAGGTGTGTCACACGCAACCACTGATGATG3CAACAAGGCTCTCAA 1582
QY 2006 CAGTGTGATGATGATGAATCATTTGCTGATTCAGGGTAAAGTATGCGCAAGTGTAT 2065
Db 1583 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1642
QY 2066 CCAACCATGTTTCTGCTTAACCTCATAGTATCTAATGAAGATGTGCAATACAGA 2125

Db 1643 GACACGAGGCTGCTGTTCTC---CCGAGGTGGGGTATGCTGACATGTGGCAGCTGACG 1699
QY 2126 GCAGCGGTGTCAGAACCATGATTAACCCAA 2156
Db 1700 AGGACAGAGAAACCTTGAAGACACTTA 1730

RESULT 3
US-09-974-300-6511
; Sequence 6511, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groch
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6511
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6511

Query Match 7.9%; Score 185.8; DB 10; Length 939;
Best Local Similarity 56.4%; Pred. No 7.4e-42;
Matches 346; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 1477 TTGAAGTCCCTGACAGAAAGAAAGTGGCCATTTGAGAGTATGAGTGTGAGGAAAGC 1536
Db 285 TTAAACATTCACAGAGGACAGACGGTTGCCCTTGTGGCATGAGCGGTGAGGGAAGC 344
QY 1537 ACAATATGAGGCTATTTATTTGCTTCTATGAGCTTCAAAAGGTAGCATTTATCTTGTCT 1596
Db 345 ACGTCTCTGAGCTGATACCGAGGTTTACATGTTTACAAAGCCGATTCGATGAC 404
QY 1597 GGTCAAAATATTAAGAATGTAGCTGGAAGCCTTGGAGGCGAGTGGAGTGAATCCT 1656
Db 405 GGCAGAAACATTTGTGACCTCTGCTGAGAGTTTGGCGCAATATTTGCAATGCTTCTG 464
QY 1657 CAGATGCTGTCTCTTCCATTAATTAATTAATTAACCTCTTATATGAGAAATCAGT 1716
Db 465 CAGATTAATATCTGTTTACGACAGCGTGAATGCAATTAATAATTTGAAAGCCTGAT 524
QY 1717 GCTTACCTGAGAGTGTATGACGTGGCAAAATTAAGTGAATTTGCAATTTCTT 1776
Db 525 GCAACAGATTAAGATCAATTTCAAGGACGCAAAAGCCCAACCCATTAATTTTGC 584
QY 1777 CGAATGCCATGATATGACACCAAGTATGGGAAACGAGCACTTCAAGCTTTCAAGGGA 1836
Db 585 AAGCTGAATATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 644
QY 1837 GAAAGCAAGATAGCAATTTGCAAGCCATTTTGAAGAACCCCAAGTCAATCTAT 1896
Db 645 CAAGAAGAGGATGCAATTTGCGCGGATTTTAAATAATCACTGTGCTCATTTT 704
QY 1897 GATGAAGTATCTCATGCTTATGATTTGATTACTGAAGAGACTATTTCTGCTGCAATGAAG 1956
Db 705 GAGGAAGCAACGTCAAGCGCTGATTTGAAAGCAATTTTATCCAGAAAGCACTAGAA 764
QY 1957 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2016
Db 765 AACTGAGCCGGAACCGGAACGTTTATGCTGACATGCTTTGTCACATTTAGCAAT 824
QY 2017 GCAGATGAATCATTTGCTTGTGATCAGGGTAAAGTATGAGCCGAACGTGTACCACTAGT 2076

Db 825 GCCGATGAGTGTGCTATATAGACGCGAAAGTAGTGAACAAGGCACTCATAAAGA 884
Qy 2077 TTGCTTGCTAAC 2089
Db 885 TTAATGGAAGG 897

RESULT 4

US-09-815-242-8403

; Sequence 8403, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8403
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1737)
US-09-815-242-8403

Query Match 7.9%; Score 184.6; DB 10; Length 1737;

Best Local Similarity 56.8%; Pred. No. 2.4e-41;
Matches 340; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

Qy 1491 AGGAAGAAGGCGCATTTAGAGGTAGTGGTCCAGGAAAGCACATATAGAGCT 1550
Db 1098 AGGAAACAGTCTTCTTGTAGTATAGTGTGTGTATCAACTTAATTAAGT 1157
Qy 1551 ATTATTTGCTTCTATGAGCCTCAAAAGGTAGCAITTAATCTTGTGTCAAATATACA 1610
Db 1158 AATACCGAGATTTAGATGTAATCTTGGCAAAATTTAATAGATGTCACAAATTA 1217
Qy 1611 AGATGTGAGCTGGAAGGCTTCGAGGAGGAGTGGAGTGTACTCAGAGATGCTGTCT 1670
Db 1218 AGATTTTAAAGGGAAGTTTAAGAAATCAATAGATTGGGCAACAGATTAATATTT 1277
Qy 1671 CTTCATATATCTATTATTAACACTTATATGGAACATCAGTCTTCACTGAGGA 1730
Db 1278 ATTCTCCGATAGTTAAGAAATATTTTACTTGTCTCCACAGCAACAGATGAGGA 1337
Qy 1731 AGTGTATGAGTGGCAAAATTTAGCTGAGACTTCATGATGCAATTCCTGCAATGCCAATGG 1790

Db 1338 AGTAGTGAAGCGCGCAAAATAGCTTAATGACATGACTTTATATGAATTGCCACAGGG 1397
Qy 1791 ATATGACACCCAGTAGGCGGAACAGAGACTCAGCTTTACAGAGAGAAAGCAAGGT 1850
Db 1398 ATATGACACTGAGTAGGTGAACAGAGTGTAAATTAATCAGGTGTCAAAAACAAAGTT 1457
Qy 1851 AGCAATTGCAAGAGCAATTTTGAAGACCCCGAGTCATACCTATATGATGAAGTACTTC 1910
Db 1458 ATGATTTGCTAGAAATATTTTAATTAATCCGCCAATCTTATCTCGATGAAGCAACAG 1517
Qy 1911 ATCGTTAGATTCGATTTACTGAAGAGACTATTTCTTGTCGCAATGAAGATGTGTCAACA 1970
Db 1518 TGCACTGATTTTGAAGAGTGAATCATTTATCAAGAGCATTAATGATGTGTGAAGTAAGA 1577
Qy 1971 CAGAACTTCATTTTCAATGACACAGATTGTCAACAGTGTGATGACATGAATCAT 2030
Db 1578 TCGAACGACCTTATCGTAGGCGATCGTTGTCTACATATTACATGTGTGACAAATGT 1637
Qy 2031 TGTCTTGATCAGGGTAAAGTAGCCGAACTGTGTACCACCATGTTGCTTGCTAAC 2089
Db 1638 CGTAATTGAATAATGGGCAATATTGTTGAAACAGTAGCATCGTAATGATTGCAAAAC 1696

RESULT 5

US-09-815-242-6522

; Sequence 6522, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6522
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1716)
US-09-815-242-6522

Query Match 7.8%; Score 182.4; DB 10; Length 1716;

Best Local Similarity 53.0%; Pred. No. 9.8e-41;
Matches 390; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

Qy 1349 TAGACACCAATTAAGACAAAGTAGTGCATCTCCCTTCAGATACACACAGACAG 1408
Db 932 TGAACAAACACCGTCATCAAAAGATTGCGAGAGCTGTGTCTTCAACATTTAGAG 991


```

Qy 1409 CTACCGTGGCTTGGTAAATGTGCAATTTGAAATCAATGAGGGCCAGAAAGCTCTAGTG 1468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 992 GAACGATGCTCTAATAAGATGTTTCATTTGCTATGAAGATGCAACAAAGTTTATGACC 1051
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1469 GAATATCCTTTGAAGTCCCTCCAGAAAGAGAGGCGCATGTGTAGAGTGTGGTCAG 1528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1052 ATATCAATTTAAATTAATTAACCTGTGTAAACCGTTGCTTGTGTGGCCAAAGTGTTCAG 1111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1529 GAAAGACACAAATAGTAGGCTATTAATTCCTCTATAGACCTCAAAAGGGTAGCATTT 1588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1112 GAABAAAGACCTTGTGCAATCTGTTGCTCTTTTAAAGTAGTAAAGTGTGAATTA 1171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1589 ATCTTGCTGTCAAAATATACAAAGTGTAGCTGTGAAAGCCTTGGAGGGCAGTGGAG 1648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1172 CGATTTAGCGGTAGAAATATCCAGCAATAGACCTTGGCTTCATTAAGTAAATTTGAA 1231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1649 TGTGACCTCAGAGTGTCTCTCTCCATATATCTATTAATTAACACCTCTAATATGAA 1708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1232 TTGTTCAACAAAGATGTGTTTATTTATCTCTGGACATTAACGAAATTAATGCTTATGGA 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1709 ACATCAGTGTCTTCACTGAGGAAGTATGCACTGGCAAAATTAAGCTGATCATGATG 1768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1292 ATTTAAATGCTACGAATATAGACATTCACAAAGCGGTGAATTTGACATTTAAGAACATG 1351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1769 CAATTTCTGCAATGCCACATGATATGACACCAAGTAGGGGAACGAGACTCAAGCTTT 1828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1352 TAAATTCAGCTTAATCCAGATGCTTAGACACAAATTAATTTGGGGAACGTGGGTGAATTA 1411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1829 CAGAGAGGAAAGCAAAAGATAGCAATTTGCAAGGCCATTTTGAAGACCCCGGAGTCA 1888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1412 CGGGCGGTCAAAAGCAGATGCAATGCGCAATGTCGCAATGTTCTTAAAGACCCCGGATTT 1471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1889 TACTCTATGATGAAGTACTTCTCATGCTTAAATGATTAAGATTAAGAGACTATTTGCTG 1948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1472 TAAATTCAGATGAAGCCATTTCTGCATTTAGATACAGAAACAGACAGTGTATTAAGAT 1531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1949 CCATGAAGATGTGTCAAAACAGAACTTCTATTTTCATTTGACACAGATTTGTCAACAG 2008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1532 CATTAACCTGTGTAGTACGGGCGAACAACATTAATTAATTTGCGCACAGCTGGCAACGA 1591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2009 TGGTGTATGCAATGAATCATTTGCTTGTGATCGGGTAAGGTACCGAACGTGTATCCC 2068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1592 TTAAACATGTGATCGAATATCGTAGTGTACCAAGAAATTTTGAAGATGGAACAC 1651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2069 ACCATGTTTGTCTGC 2084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1652 ATGAAACTTTGTACGC 1667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-09-956-004-79/c
; Sequence 79, Application US/09956004
; Patent No. US20020072595A1
; GENERAL INFORMATION:
; APPLICANT: Patrick J. Dillon et al.
; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
; FILE REFERENCE: PB324D1
; CURRENT APPLICATION NUMBER: US/09/956,004
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 08/976,259
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/061,953
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: 60/031,626
; PRIOR FILING DATE: 1996-11-22
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 3576
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURES:

```

```

; NAME/KEY: misc.feature
; LOCATION: (1528)..(1528)
; OTHER INFORMATION: n equals a, t, g, or c
; NAME/KEY: misc.feature
; LOCATION: (2618)..(2618)
; OTHER INFORMATION: n equals a, t, g, or c
US-09-956-004-79

```

```

Query Match      7.1%; Score 165.4; DB 10; Length 3576;
Best Local Similarity 53.7%; Pred. No. 9.1e-36;
Matches 340; Conservative 1; Mismatches 292; Indels 0; Gaps 0;

```

```

Qy 1492 GAAAGAAAGTGGCCATTTGAGAGGTAGTGGTCAAGGAAACAAATAGTAGAGCTTA 1551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3564 GGGAGAGTTATTTGATTTGTTCGAGCGTTCTGTTACAGAAAGACATTAATTAATTA 3505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1552 TTATTTGCTTCTATAGCTCAAAAGGTAGATTTATCTTGTCTGTGTAATATACAA 1611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3504 ATTCAACGTTTTATATTTCTTGAAATGCGCAGGCTTTAATGTATGACATGATCTTGGC 3445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1612 GATGTAGCCTTGAAAGCCTTCGAGGAGCAGTGGAGTGTGTAACGATGCTGCTTC 1671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3444 TTGGCCGATCTTAACGTGTGTTACGTCTCAGGTGGGGTGTGTGACAGACATGTGCTG 3385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1672 TTCCATTAATCTATTTATTAACAACCTTTATATGAAACATCACTTCACTGAGAA 1731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3384 CTTAATTCGAGTATTTATGTAATTAATTTCACTGCTAATCTCGCATGTCGTCGAAAAA 3325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1732 GTGTATGCACTGGCAAAATTAAGTGTGATTCATGATGCAATTTCTTGAAATGCCATGGA 1791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3324 GTTATTTATGACAGCAAAATTAASCAAGTGTCTCATGATTTTATTTCTAATTTGCGTGAAGGGG 3265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1792 TATGACACCCCAATAGAGGGAACAGAGACTCAAGCTTTCAGAGAGAAAGCAAAAGATA 1851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3264 TATAACCACTTTGTGGGGAACAGGGGCAAGATTTTCGGAGGTCAACCTCAACGCATC 3205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1852 GCAATTCGAAGACCAATTTTGAAGACCCCGCATCATCTATGATGAGTAAAGTACTTCA 1911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3204 GCAATTCGAAGGCGCTGTGTGAACCACTTAATACTCATCTTTGATGAAAGCAACAGT 3145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1912 TCGTGTATTCATTTCTGAAGACTATTTCTGTGTGCAATGAAGATGTGTCAAAACAC 1971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3144 GCTGTGATTTATGATGTGAGCAATGTCATATGCGAATATGCAAAATATGTAAAGGC 3085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1972 AGAATCTTATTTTCATTTGACACAGATGTCAACAGTGGTGTGATGCAATGAAATCATTT 2031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3084 AGAACGTTTATATCACTTGTCTATGCTGTCTTACAGTAAAAAATCAGACCCCATTAAT 3025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2032 GTCTTGATCAGGTTAAGGTAGCCGACAGTGTATCCACCATGTTTGTCTTAAACCTT 2091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3024 GTCATGAAAGGGAAGAAATTTGTTGAACAGGGTAAACATTAAGAGCTGCTTTCTGAACCG 2965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2092 CATAGTATCTATTCAGAAATGTGCAATACACAG 2124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2964 GAAAGTTTATACGTTACTTATATACATTACAG 2932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-09-974-300-2176
; Sequence 2176, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27

```

```
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2176
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2176
```

```
Query Match      6.7%  Score 158.2; DB 10; Length 1233;
Best Local Similarity 53.9%; Pred. No. 4.8e-34;
Matches 325; Conservative 0; Mismatches 278; Indels 0; Gaps 0;
```

```
QY 1472 TATCTTTGAAGTCCCTGAGAGAAAGAAAGTGGCCATTGTAGAGAGTGGTCCAGGGA 1531
DB 576 TTTCTTTGAAGCCGACGAGGCGAAAGCGTGGCGCTGCGGATGAGCGGAGCGGGA 635
QY 1532 AAAGCAATAGTAGGCTATTATTTCGCTTCTATGAGCCTCAAAAAGGTAGCATTTATC 1591
DB 636 AATCAGCGCTGTCCAGCTTCAGAGATCTCAGCATGCTCAGCCGAGACCTGGAGA 695
QY 1592 TTGCTGTCAAAATATACAAAGATGAGCCTGGAAAAGCCTTGGAGGGGAGTGGAGTGG 1651
DB 696 TCGAGCGATCGATGTCGCGGATATAGAGCGCGAAAGCCTTAGAAACCAATCGCATGG 755
QY 1652 TACCTGAGATGCTGTCTCTTCATTAATATCTATTATTACAACTCTTATATGAAACA 1711
DB 756 TTCTCCAGGATACGTTTTATTACAGGATCGGTAGCGGAAACATATCGATCGGGAATC 815
QY 1712 TCAGTGTTCACCTGAGAGAGTGTATGCAAGTGGCAAAATTAGCTGACCTTCATATGCA 1771
DB 816 CCGAAGCTTCAGTGAAGACATCATTTGAAGCGGCAAAAGCTCGAATGCCATGATTTTA 875
QY 1772 TTCTTCGAATGCGCATGATATGACACCCAGTAGGGGAGAGAGAGTCAAGCTTTTCAG 1831
DB 876 TTATGAGAGTCCGCGAAGGCTATGATACAAAGTTCGAGAGAAAGGAGTCAAGCTGTCCG 935
QY 1832 GAGGAAAGCAAGAGATGCAATTTGCAAGAGCCATTTTGAAGAACCCCGCATATAC 1891
DB 936 GAGGCAAAACAGCGGATATCATTCGCGAGGGTGTCTTGAAAGAACCGCGCTTTAA 995
QY 1892 TCTATGATGAGACTTCTATGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1951
DB 996 TTTTGAATGAAAGCAGTCAAGCTTGAAGCTTGAAGAGCAGCATATATCCAGGAAGGA 1055
QY 1952 TGAAGGATGCTGCAAAACAGACCTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 2011
DB 1056 TGGAAAGCTCCGAAAGACCGGAGACCTTTATGTTGCTCAAGACTGTGACGATTA 1115
QY 2012 TTGATGCAAGTAAATCATTTGTTGATCAGGTAAAGGTAGCCGAGCTGTAGCCACC 2071
DB 1116 CCCAGCTGACAAATCGTTGTTCATCGAAGAAACGAGCGTGTGCGGAAATCGGCACACATG 1175
QY 2072 ATG 2074
DB 1176 ACG 1178
```

```
RESULT 8
US-09-070-927A-275/c
; Sequence 275; Application US/09070927A
; Patent No. US2002012016A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunach
; Patrick J. Dillon
; Steven Barish
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
```

```
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 275:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 275:
US-09-070-927A-275
```

```
Query Match      6.7%  Score 157.8; DB 10; Length 6415;
Best Local Similarity 53.5%; Pred. No. 1.8e-33;
Matches 327; Conservative 1; Mismatches 283; Indels 0; Gaps 0;
```

```
QY 1477 TTTGAAGTCCCTGAGAGAAAGAAAGTGGCCATTGTAGAGAGTGGTCCAGGAAAGC 1536
DB 6041 TTTTCGTTGCAAGAGAGAAACCTTGGGATTTTGGCAAAACAGGCGCTGTAAC 5982
QY 1537 ACAATAGTAGGCTATTATTTCGCTTCTATGAGCCTCAAAAGGTAGCATTTCTTGCT 1596
DB 5981 ACTATTTGAAATTTGATCGTGAATATGACCAATATCAAGGAACGATTTCTTTGGA 5922
QY 1597 GGTGAATATATCAAGATGAGCCTTGGAAAGCCTTGGAGGGAGTGGAGTGTACT 1656
DB 5921 AAACATTAACATTAATAATTAACATTTGATGATGATGATGATGATGATGATGATGATG 5862
QY 1657 CAGATGCTGTCTCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
DB 5861 CAGATCAATTTCTCTTTTCATGACGATGACGATGACGATGACGATGACGATGACGATG 5802
QY 1717 GCTTACCTGAGAGTGTATGACGTGGCAAAATTAGCTGACTTCATGATGATTTCTT 1776
DB 5801 TTGAAACAGAGAGATGAAACAGCGGAGCATTTAGCATTTATTAACAAAGAAATTAA 5742
QY 1777 CGAATCCACATGATATGACCAAGTAGGGGAGAGAGAGCTCAAGCTTTCAGAGGA 1836
DB 5741 GCATTCCTGAGGCTATGACAAATGTTGGGAAAGTGGCGTTTGCCTTTTCAGAGGA 5682
QY 1837 GAAAGCAAGAGTAGAATTTGCAAGGCCATTTTGAAGAGCCCGGAGTCACTAT 1896
DB 5681 CMAAAACAGAGATTTCTATTGACAGGGGCTTGATTTGTAACAGAAATGTTAATTTTA 5622
QY 1897 GATGAGTACTTATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1956
DB 5621 GACGATGCTTATCTGAGTGTGATGCAAGACTGAAGAGGCGATTTATCAATCTAA 5562
QY 1957 GATGTGCTAAACAGAGATTTCTATTTCATTTGACACAGATTTGTCAACAGTGTGAT 2016
DB 5561 GAAAGCGGCAAGAAACCACTATTATTATACAGCGATGTTAAGTAGTGTATCAT 5502
```

Qy	2017	GCAGTGAATCATTCCTTGGATCAGGGTAAAGGTACCGAAACGGGTATCCACCATGT	2076
Db	5501	GCCTAAAGAACTTAGCTTGGATGAAGGAAATATTATTGAACGGGGACGATCTCAAG	5442
Qy	2077	TTGCTTGCTAA	2087
Db	5441	TTACTGGCTCA	5431

RESULT 9

```

US-09-880-107-2299
Sequence 2299, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2299
LENGTH: 3924
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 M32334
US-09-880-107-2299

```

Query Match	6.7%	Score 157.4	DB 10	Length 3924
Best Local Similarity	53.5%	Pred. No. 1.7e-33		
Matches 329	Conservative	0	Mismatches 286	Indels 0
				Gaps 0

Qy	1456	AAAGTCCTTAGTGGAAATATCTTTGAAGTCCGCGAGGAAAGAAAGTGGCAATGTGTAGA	1515
Db	1260	AMAGTCTTGAAGGGCCTCAACCTGGAAGGTGCAGATGTGGCACAACGGTGGCCCTGTGTGGA	1319
Qy	1516	GGTAGTGGGTCAAGGAAAAGACAATAGTAGGCTATATTTCCGCTTCTATGAGCTCAA	1575
Db	1320	AGTAGTGGCTGTGGGAAAGACACAACGGTCCAGCTGATACAGAGGCTCTATGACCCTGAT	1379
Qy	1576	AAGGCTAGCATTTATCTGTGCTCAAAATATACAGAATGTAGGCTGGAAAGCTTCGG	1635
Db	1380	GAGGGCAAAATTAACATGTATGGCGAGAAATATAGAACCTTTAATGTAACTATCTGAGG	1439
Qy	1636	AGGGCAGTGGAGTGTGTAACCTCAGATGTGCTCTCTTCATATACTATTTATTACAC	1695
Db	1440	GAATATCTGTGTGGTAGTACAGAGCCGGTGTCTGTTTCCACCAATGTCTGAATAAT	1499
Qy	1696	CTCTTAATAGGAAACATCACTGCTTCACTGAGAAAGTATAGCAGTGGCAAAATTAGCT	1755
Db	1500	ATTTGTTATGCGCGTGGAAATGTAAACATGATGTAGATATAAAGAAAGCTGTCAAGAGGCC	1559
Qy	1756	GGACTTCATGATGCAATTCCTTCAATGCCACATGTATGACCCCAAGTAGGGGAAAGCA	1815
Db	1560	AAGCGCATGTAGTTTATCATGAAATTAACAAGAAATTTGACACCCTGTGGTGTGAGAGAGA	1619
Qy	1816	GGACTCAAGCTTTACAGAGAGAGAAAGCAAAGGTAGCAATTTGCAAGGCCATTTTGAAG	1875
Db	1620	GGGGCCACAGCTGAGTGTGTGGCGAAGACAGAGAAATCGCATTTGACGTCGGCCCTGATTGCG	1679
Qy	1876	GACCCCCAGCTCACTACTATGATGGAAGTACTTCACTGTTATGATTTGATTACTAGGAAG	1935
Db	1680	AAACCCCAAGATCTTCTGTCTGATGTAGGCCACGTACGCAATTGTGACACAGAAAGTBAAGCT	1739
Qy	1936	ACTATTTCTGTGCCATGAAAGATGTGTCAACAACAGAACTTCTATTTTCAATTGACAC	1995

Db 1740 GAGGTACAGGAGCTCGATTAAGGCAAGAGCGGACCAACCATTTGTGATTACACAC 1799

QY 1996 AGATTGTCCACAGTGGTTGATGCAGATGAATCATTTGTTGGATCAGGGTAAGGTAGCC 2055

Db 1800 CGATGCTCTACGGCCGAATGCAGATGTCATCGCTGGTTGAGAGATGAGACTAATTGTG 1853

QY 2056 GAACGTGTACCCAC 2070

Db 1860 GAGCAAGGAAGCCAC 1874

RESULT 10

US-09-070-927A-255/c
Sequence 255, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunesh
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 255:
SEQUENCE CHARACTERISTICS:
LENGTH: 3601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 255:
US-09-070-927A-255

Query Match	6.6%	Score 155	DB 10	Length 3601
Best Local Similarity	52.8%	Pred. No. 7.5e-33		
Matches 335	Conservative 0	Mismatches 300	Indels 0	Gaps 0

QY 1490 CAGGAAAGAAAGGCCATTGTAGAGAGTAGAGGCTCAGGGGAAAGCAATATGAGAGGC 1549

Db 1716 CAGGAGAAACAATGCCATTGTTGGCGCACTGTGCGGGTAAATCAACATTAGCAATT 1657

QY 1550 TATTTATTTGCTTCTATGAGCCTCAAAAGGTAGCAATTATCTTGCTGCTCAAAATATAC 1609

Db 1656 TGTTTAATGCGCCTTTTACGAGGTGACCGTGGGCAAAATTACGGTTATATCTCACGCACTTA 1597

Qy	1610	AAATGTGAGCTGGAAAAGCTTGTGAGAGGCAAGTGGAGTGAACCTGACAGATGCTGTC	1669
Db	1586	CTCAAAATTAATCTCGGGACAGCTTGGGAAAAAGTTTGTATGGTCTCCAAAGTACTTGCC	1537
Qy	1670	TCCTTCATAATACTATTATTATTACACCTCTTATATGGAACATCATGTCCTCACCTGAG	1729
Db	1536	TATTCGATAGCTCATCTCGGAAAAATTAAACGTACGGAATCTGAAGCTCTGACGAG	1477
Qy	1730	AAGGTATGACGTGGCAAATTTAGCTGACATTCATGATGCAATTTCTTGAAATGCCACAG	1789
Db	1476	AAATTAATGAAAGCCATGAAAAAGGCCATATTTTGTATTTGGTCATGCGTTTACTCTCAG	1417
Qy	1790	GATATGACACCCCAAGTAGGGGAAACGAGACTCAAGCTTTCAGAGAGAGAAAAGCAAAG	1849
Db	1416	GACTAGATACCGAATATGGTTTCCCAAGGCGTTAAATTTCTGAAGACACACACATTTAA	1357
Qy	1850	TAGCAATTTGCAAGGCCATTTTGAAGACCCCCCACTCATCTCTATGATGAAGCTATT	1909
Db	1356	TGACTATTGACACGACCATATTAATGATATCCGAGTCTTATTTTGAATGAAGGACCA	1297
Qy	1910	CATCGTTAGATTCGATTACTGAAGAGACTTTCTTGTCCTCATGAAGATGTGTCAAC	1969
Db	1296	GTTCAAGTTAGACATGACCGAGGAAAAAATTCAGATGCTTTTTTAACAATGATGACTG	1237
Qy	1970	ACAGAACTTCTATTTTTCAATTGACACAGATTTGCAACAGTGGTGTATGCAATGAATCA	2029
Db	1236	GCCGACCACTTTTGTATTGTCCTCATCGCTTACCAACATCAAAAGTGTGAAAAATTC	1177
Qy	2030	TTGTCTTGATCAGGTGAAGTGAAGCCGAACGTGTATCCACCAATGGTTGTGCTAAC	2089
Db	1176	TAGTATATGATATAGTCATCAAGTGTGAATTGGCAATCCATATGCAATATTGCAAAAAG	1117
Qy	2090	CTCATAGTATCTATTCAGAAATGTGGCATACAG	2124
Db	1116	AACAGGGTATTATCTGACCTTATCGAACACAG	1082

```

RESULT 11
US-09-974-300-2172
Sequence 2172, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
TITLE OF INVENTION: Expression
FILE REFERENCE: 10085, 500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,558
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq For Windows Version 4.0
SEQ ID NO 2172
LENGTH: 1791
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-2172

```

Query Match	6.6%	Score 154.6	DB 10	Length 1791
Best Local Similarity	50.9%	Pred. No. 6.2e-33		
Matches 367; Conservative	0	Mismatches 354	Indels 0	Gaps 0

Oy	1371	AGTATGSCATCTCCCTTGAGATCACACAGACAGCTACCGGCGCTTGTAAATGT	1430
Db	1037	AGACAGGAGATCGGCTCAAAATTAAGAATTAAAGCCGAGTCGATTTGAACAGGT	10986
Oy	1431	GCATTTGATATCATTAAGGCGCAAAAGTCTTAAGTGATATCTCTTGAAGTCCCTGC	1480
Db	1097	GGAATTTCTTATATACGGAAACCGCAAGGCGCTGCACAGCATTTCTTTGAATATCTGCG	1156

Qy	1491	AGGAAGAAAGTGGCCATTGTTAGAGAGTGTGGGTCCAGGAAAAGCAAAATGAGAGCT	1550
Db	1157	GGGATCATCGCTTGGCTTGTGTGGACATACAGGAATCAGGGAAGACAGATCGCAATCT	1216
Qy	1551	ATTATTTCCGCTTCTATAGAGCTCAAAAGGGTAGATTATCTTGTGTGCTCAAAATATACA	1610
Db	1217	GATCAGCAGGTTTTATGAGTAGCAACAGCGGGGCTGTCAAAAGTAGACGGTATACGATATAC	1276
Qy	1611	AGATGTAGGCTTGGAAAAGCTTTCCGAGGGCATTGGGAGTGTGTAAGTCTGATCTCT	1670
Db	1277	CGAATATCTGCTTTCCGATCTTCGTTGTGGAAATCAGCAATCGTGTCGACGGACAGTTTAT	1336
Qy	1671	CTTCGATAAATACTATTATTTATACACCTCTTATATAGGAAACATCAGTGCCTTCACTGAGA	1730
Db	1337	TTTTTCCGGAACGATTTATGGAATAATTTGTTTGGCCGTCCGAACGCCGCATGTAAGA	1396
Qy	1731	AGGTATATGACGTGGCAAAATTAGCTGGACTTATGATGCAATTTCTTCSAATGCCATCG	1790
Db	1397	AGTATCGAAGCAGCAAAAGCTGTCCGAGCCGATTCGTTTATTCAGCGCTTCCGAATGG	1456
Qy	1791	ATATGACACCCAGTAGTAGGGGAACGAGSACTCAAGCTTTACGAGGAGAAAGCAAGAGT	1850
Db	1457	ATAGCCACAGAAAGTGAAGAAAGCGGCAATGATTTATGCGAGGTGAGCGCAAGCTAT	1516
Qy	1851	AGCAATTCGCAAGAGCATTTTTGAAGAGCCCCAGCTCATCTTATGATGTAAGTACTTTC	1910
Db	1517	TTTCAATTTGCCAGGGCCCTTCTGCGAGATCTCTGATTTCTAATCTTGATGAAGCAACGC	1576
Qy	1911	ATCGTAGATTCGATTACTGAAGAGCTATTTCTTGGTGCATGAAGATGTGTCAACA	1970
Db	1577	AAGCATTTGACACGAAAGCGGAAGTCAAAATCCAGGCTCCCTTGAAMACATCGCTCTACGG	1636
Qy	1971	CAGAACCTCTAATTTTATTTGACACAGATTTGTCACAGTGGTGTATGTCAGATGAAATCAT	2030
Db	1637	CAGAACGGCTATTATATATGCTCTCACAGGCTTTCAACGATAGGATATCGGATTAATAATCT	1696
Qy	2031	TGTTCTTGATCAGGGTGAAGGTAGCCGAAAGTGTATCCACCAATGCTTTGCTGTAAACC	2090
Db	1697	GGTTCTTGAACAGCGCCGGMAAATAGAAAGAAATATCCAGAAATTTGATCGGGAAGAC	1756
Qy	2091	T 2091	
Db	1757	T 1757	

RESULT 12
 US-09-070-927A-12/c
 ; Sequence 12, Application US/09070927A
 ; Patent No. US20020120116A1
 ; GENERAL INFORMATION:
 APPLICANT: Charles A. Kunisch
 Patrick J. Dillon
 Steven Barash
 TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
 NUMBER OF SEQUENCES: 982
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/070,927A
 FILING DATE: 04-MAY-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16
 APPLICATION NUMBER: 60/04,031
 FILING DATE: 1997-05-06
 APPLICATION NUMBER: 60/066,009
 FILING DATE: 1997-11-14
 ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302
 REFERENCE/DOCKET NUMBER: PB369
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9047 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-09-070-927A-12

Query Match 6.6%; Score 154; DB 10; Length 9047;
 Best Local Similarity 47.9%; Pred. No. 2.6e-32;
 Matches 505; Conservative 0; Mismatches 545; Indels 4; Gaps 2;

QY 1036 TTATAATGAAGATGAAGACAGATATGATGATTTTGAAGACGTATGAGACT 1095
 DB 8778 TTCAATATCAAGATTTTGAAGACAGCAATCAATTAATAATATG-CTGA 8720
 QY 1096 GCTTCATTGAAAGTACTCTACTCTGCTATGCTGAACTTTGGTCAAGTCTATTTTC 1155
 DB 8719 ACGCATGATGAAGTGAATCAAACTCTTGCTGATTAATCTCCGCTGCTTTTATATG 8660
 QY 1156 AGTGCCTTAAACAGCTATATGATGCTGCGCAGTCAAGGAAATTTGGAGTACCTT 1215
 DB 8659 GGCCTTGTACGCTCTATTTTGTGCTTGGCGTATTTAGTTGACATGAGCACTTA 8600
 QY 1216 ACTGTGAGATCTAGTAATGATGATGACTGCTTTTTCAGCTTTCATTAACCCCTGAC 1275
 DB 8599 CAATTTGGACCGTTACAGCTGTGATGATATAGCGTTCACTTGGCTTATATG 8540
 QY 1276 TTTCTGGAATCTATATAGAGACTAGACACTCATATATGAACTTTGTT 1335
 DB 8539 ATCCGGCTAGGATTAATGATGATTCACAGGTCCTGCTTCAATCGCTTTCAA 8480
 QY 1336 ACTCTACTCAAGTAGACACCCCAATTAAGCAAAAGTATGATGATCCCTCAATC 1395
 DB 8479 GAACTTTTGTACAGAAATTTGAATTAAGCATCTCATTAAGGCAACCATTTGTTAT 8420
 QY 1396 ACACACAGACAGCTACCGTGGCTTTGATATGATGATTTTGAATCAATTGAGGCGCAG 1455
 DB 8419 CATCTGAGAAAGCTTATGCTTGCATGACGTCAAGTTTCAATACAGAAACAGCT 8360
 QY 1456 AAGTCTCTTATGAGAA--TATCTTTGAAGTCCCTGAGAGAAAGAAAGTGGCATTGTA 1512
 DB 8359 GATCTCTTTTGAAGAAATGATGATTTGTCAATCTTAAGGAAACCAACGCGCATTTGT 8300
 QY 1513 GGAGTAGTGGGTACAGGAAAGCAAAATGATGAGCATTAATTTGCTTCTATGAGCCT 1572
 DB 8299 GGTCAACTGGCGCTGTAAGAAAGTACTTATTAAGTTACTTTTCAATTAAGAGGTC 8240
 QY 1573 CAAGAAGTATGATTTTCTGCTGTCAAAATATACAAAGTATGAGCTGGAAGCCTT 1632
 DB 8239 ACAGCCGACAGATTAGCTATTTGCGACAGATATCCGCTCATTAATCTCAGCAACAGATT 8180
 QY 1633 CGAGGCAAGTGGAGTGTACTTCAGAGTGTCTCTCTTCCATTAATACTTATTTATAC 1692
 DB 8179 CGCCAAAGTCAATGATTAATGTCACAAAGACCTTTCTTTTCAATGAGCAATCTTATCA 8120
 QY 1693 AACCTTATATGGAATCATCAGTCTTCACTGAGAGATGATGAGTGGCAAAATTA 1752
 DB 8119 AACTTATTAATGGAAATGCCAAAGCAATACAGAAAGTAATAGACGCGCTATGAATTT 8060

QY 1753 GCTGACTCATGATGCAATCTTTCGANTGCCACATGATATGACACCCAGTAGGGAA 1812
 DB 8059 TCACAAATCTTCTGAATTTATTCGATTCCTTACCAAGGAGTGAAGATTTCGTAGCACA 8000
 QY 1813 CGAGGATCAAGCTTTCAGAGAGAAAGCAAGAGTACATTTGCAAGAGCATTTTG 1872
 DB 7999 GGGGGTCCAACTATTTCTGTTGTCAAAACAAAGATGTGATTTGACAGGCTTATTC 7940
 QY 1873 AAGGACCCCGACGCTACTCTATGATGAACTACTCATCTGTTAGATTGATTAAGTAA 1932
 DB 7939 AAACCGGACAGCTTATATTTTCGATGATGATTTTCCGCAATTAAGTTACAAACGAT 7880
 QY 1933 GAGACTATTTCTGTTGCTGATGAAAGATGTGCTAAACACAGAACTTATTTTCAATGCA 1992
 DB 7879 GCGCTCTACGTCGCTGCTTACATGACAAATGTCCGACAAACTTACTCATATTTGCT 7820
 QY 1993 CACAGATTGTCAACAGTGTATGATGAGAAATATCTGTTGATTCAGGGTAAAGTGA 2052
 DB 7819 CAACGTTAAAGTCAATCAATGAAACGCTGACACATTAATGCTTCAATGAAAGGATTT 7760
 QY 2053 GCCGAACGTGTACCCACCATGCTGTTGCTGCTA 2086
 DB 7759 GTTGTCAAGGACCCACGCTGATTTACTTACCA 7726

RESULT 13
 US-09-938-842A-2237
 Sequence 2237, Application US/09938842A
 Patent No. US20020160378A1
 GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT: Kreps, Joel
 APPLICANT: Wang, Xun
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 TITLE OF INVENTION: SAME, AND METHODS OF USE
 FILE REFERENCE: SRIPI300-3
 CURRENT APPLICATION NUMBER: US/09/938,842A
 PRIOR FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: US 60/227,866
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: US 60/264,647
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/300,111
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 5379
 SEQ ID NO 2237
 LENGTH: 3861
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2237

Query Match 6.5%; Score 153.2; DB 9; Length 3861;
 Best Local Similarity 51.8%; Pred. No. 2.5e-32;
 Matches 347; Conservative 0; Mismatches 323; Indels 0; Gaps 0;

QY 1454 AGAAGCTCTAGTGAATATCTTTGAAGTCCCTGAGAGAAAGAGGCGCATTTAG 1513
 DB 1196 AGCAAAATTTCCGGAATCTCTCTTTATCTCGAGTGTACAAAGTGGCTTTAGTTG 1255
 QY 1514 GAGTAGTGGGTACAGGAAAGCAATAGTAGGCTATTTATTTCCCTTATAGAGCTTC 1573
 DB 1256 GAAAGTGGAGCGGGAATCTACTGTTGTGATGTGATGAGAGTTTATGAGATCCAC 1315
 QY 1574 AAAAGGTACATTTATCTTCTGTCGTCAAAATATACAGATGTGAGCTTGAAGCCTTC 1633
 DB 1316 AAGCTGTGACGTTCTCATAGATGATTAATTAAGATTAAGATTAAGATTA 1375
 QY 1634 GAGAGGCAAGTGGAGTGTACCTCAGAGTGTCTCTCTTCCATTAATTAATTTATTA 1693
 DB 1376 GAGCAAGATGAGCTTTGATGATCAAGAACCAATTTTGTCTACTGCAAGATCAAGATA 1435
 QY 1694 ACCTTATATGGAATCATCAGTGTCTTCACTGAGAGATGATGAGTGGCAAAATTA 1753

Db	1436	ACATTGCGTACGGCAAAAGAACGCCAACACCGAAGAGATTAAAGCAGCTGCAGAGCTTAG	1495
Qy	1754	CTGACATTCATGATGCAATTCCTTGGATGCGACATGATATGACACCCAAAGTAGGGGAAC	1813
Db	1496	CAAAOGCATCTAAATTGTTGGATTAGCTATACCAAGGGTTTGGATTCAATGGTTGGAGAAC	1555
Qy	1814	GAGACATCAGCTTTTCAGAGAGGAAAAAGCAAAAGATAGCAATTCGAAAGAGCCATTTGA	1873
Db	1556	ATGGTACTCAGCTTTCCGGTGGGACAAAAAACAGAGAAATCCGGGGGCTAGAGCAATCTTAA	1615
Qy	1874	AGGACCCCCAGTACTATCTATAGTAGAAGCTACTTCATTCGTAGATTGATTAAGG	1933
Db	1616	AAGATCCAAAGATTTTACTTTTATGATGAACGTCAAGCGCGCTTATGAGATTTGGA	1675
Qy	1934	AGACTATTTCTTGGTGGCCATGAAAGATGTGGTCAAAACAGAACTTCATTTTCATTTGAC	1993
Db	1676	GAGTGGTTCAAGAAAGCCCTTGATAGAAATTATGTGTAAACCGGACTACTGTGGTGGCGCTC	1735
Qy	1994	ACAGATTGTCAACAGTGGTTGATGCGATGAATTCATTGCTTGGATCAGGGTAAAGTAG	2053
Db	1736	ATCGGTTAAGCAGCTGTGCGAAATGGCGGATATGATTCGCTGTGATCCCAAGGCAAGATCG	1795
Qy	2054	CCGAACTGGTATCCCAACCATGAGTTTGGTTCCTTAACCTCATATGATCTATTATGAAGATGT	2113
Db	1796	TGCGAAGAGGTTCTCAACGGAACCTACTAAAGAGCCGGAAAGAGAGCTATTCTCAGCTGA	1855
Qy	2114	GGCATACACA	2123
Db	1856	TTTCGTTTACA	1865

RESULT 14

US-09-938-842A-263
; Sequence 263, Application US/09938842A

```

# Patent No. US20020160378A1
# GENERAL INFORMATION:
# APPLICANT: Harper, Jeff
# APPLICANT: Kreps, Joel
# APPLICANT: Wang, Xun
# APPLICANT: Zhu, Tong
# TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
# TITLE OF INVENTION: SAME, AND METHODS OF USE
# FILE REFERENCE: SCRIPI300-3
# CURRENT APPLICATION NUMBER: US/09/538, 842A
# CURRENT FILING DATE: 2001-08-24
# PRIOR APPLICATION NUMBER: US 60/227, 866
# PRIOR FILING DATE: 2000-08-24
# PRIOR APPLICATION NUMBER: US 60/264, 647
# PRIOR FILING DATE: 2001-01-16
# PRIOR APPLICATION NUMBER: US 60/300, 111
# PRIOR FILING DATE: 2001-06-22
# NUMBER OF SEQ ID NOS: 5379
# SEQ ID NO 263
# LENGTH: 3861
# TYPE: DNA
# ORGANISM: Arabidopsis thaliana
# US-09-938-842A-263

```

Query Match	6.5%	Score 152	DB 9	Length 3861
Best Local Similarity	52.5%	Pred. No. 5.4e-32		
Matches 332	Conservative	0	Mismatches 300	Indels 0
				Gaps 0

Qy	1456	AAAGTCCTTACTGGAATATCCTTTGAAAGTCCCGACGGAAAGAAAGTGCCCATTTGTAGA	1515
Db	1150	AAGAATCTTAACAACCTTGCTCTCTGTCTCTGCTGGGAAGCAATGCTTTGGTGGG	1209
Qy	1516	GGTAGTGGGTACGGAAAGACATATGATAGGCTATTAATTTGCTTCATAGAGCTCAA	1575
Db	1210	AGCAGTGGTTACGAAAAAAGACAGTGGTTTGCCTTATAGAGGGTTTACGACCCGAAC	1269
Qy	1576	AAGGTCACATTTATCTTGCTGGTCAAAATATATCAAGATGTGAGCTGGAAGCCTTCGG	1635

Db	1270	TCAGACAAGATTACTAGATGGGCAAGACGTGAAGACGGCTAAACTGATGGTTAAG	1329
Qy	1636	AGGGCAGTGGGAGTGGTACTCAGAGTGTGTCTCTTCATTAATATCTATTATTACAC	1635
Db	1330	CAACAGATAGGGCTAGAGTCTCAAGAACTGTGATTTATTCGCACCTTCTATCAAGAGAT	1389
Qy	1696	CTCTTATATGAAACATCAGTGGCTTCACTCGAGGAAGTATGCAAGTGGCAAAATTAGCT	1755
Db	1390	ATTACTTTAGGCCGTCCGTGACGAGATCAAGTCGAGATAGAGAGAGCGTCTCGAGTGGCA	1449
Qy	1756	GGACTTCATGATGCATTTCTTGCATGCCACATGGAATATGACATCCCAAGTATGGGGAAACA	1815
Db	1450	AATGCTATTCTCTTTATCAACAACTGATGAGCTTCGACACAGAGTTGGGGAAGAGA	1509
Qy	1816	GGACTCAAGCTTTTCAGAGAGAAAGCAAAAGTAGCATTTGCAAGAGCCATTTTGAAG	1875
Db	1510	GGATTGCAGCTGTCTGGTGGGGCAAGACAAAGATAGCCATAGCAAGAGCATGTTGAAG	1569
Qy	1876	GACCCCCAGTCACTCTATGATGAAGCTACTTCATGTTAGATTGATTACTGAAGAG	1935
Db	1570	AACCGCGCATCTTCTATTAAGATGAGGAGCAGAGTGCATGGATTCTGAATCAGAGAG	1629
Qy	1936	ACTATTTCTGGTGCATGAAGAGATGAGTCAACACAGAACTTCATTTTCAATTGCACAC	1995
Db	1630	CTTGTAACAAGGCTTTGGATCGTTTCATGATTTGGAAGACATCTCTATATTTGCTCAT	1699
Qy	1996	AGATTGTCAACAGTGGTGAATGCAATGAAATCATTTGCTTGGATCAGGGTAAAGTACC	2055
Db	1690	CGCCTTCTACATTCGAAAGAGCTGACTTGTGTGCTGCTCAACAAGAAAGTGTCTCT	1749
Qy	2056	GAACGTGTACCAACAGTGGTTGGCTTGGCTAA	2087
Db	1750	GAGATTGAAGCAGATGATGAGCTTTTTCCTAA	1781

RESULT 15

US-09-738-626-2520
; Sequence 2520, Application US/09738626
Publication No. US20020197605A1

```

1  PUBLIC INFORMATION: us20020379002a1
2  GENERAL INFORMATION:
3  APPLICANT: NAKAGAWA, SATOSHI
4  APPLICANT: MIZOGUCHI, HIROSHI
5  APPLICANT: ANDO, SEIKO
6  APPLICANT: HAYASHI, MIKIRO
7  APPLICANT: OCHIAI, KEIKO
8  APPLICANT: YOKOI, HARUHIKO
9  APPLICANT: TATEISHI, NAOKO
10 APPLICANT: SENOH, AKIHRO
11 APPLICANT: IKEDA, MASATO
12 APPLICANT: OZAKI, AKIO
13 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
14 FILE REFERENCE: 249-125
15 CURRENT FILING DATE: 2000-12-18
16 PRIOR APPLICATION NUMBER: JP 99/377484
17 PRIOR FILING DATE: 1999-12-16
18 PRIOR APPLICATION NUMBER: JP 00/159162
19 PRIOR FILING DATE: 2000-04-07
20 PRIOR APPLICATION NUMBER: JP 00/280988
21 PRIOR FILING DATE: 2000-08-03
22 NUMBER OF SEQ ID NOS: 7059
23 SOFTWARE: PatentIn ver. 3.0
24 SEQ ID NO 2520
25 LENGTH: 1863
26 TYPE: DNA
27 ORGANISM: Corynebacterium glutamicum
28 US-09-738-626-2520

```

Query Match	6.5%	Score 151.8;	DB 9;	Length 1863;
Best Local Similarity	52.3%;	Pred.No. 3.9e-32;		
Matches 336;	Conservative 0;	Mismatches 307;	Indels 0;	Gaps 0;
QY	1415	TGGCCCTTGATATGTCGATTTTGATACATTGAGGGCCAGAAACTCCTTACTGGAATAT	1474	

```
Db 1079 TGAATTCAAAAGGTCACTTCGCTACGAAGAGCAAGCCGGTTATTTCCGACGTGT 1138
QY 1475 CTTTGAAGTCCCTGCAGAGAAAGAAAGTGGCCATTTGTAGAGGTAGTGGGT CAGGGAAAA 1534
Db 1139 CCATTACCGCCCGCACGCGACGCGATCGCGTGTGTGTAATCCGCGCGGTAAAT 1198
QY 1535 GCACAAATAGTGAAGCTATTTTGGCTTCTATGAGCTCAAAAGGGTAGATTTATCTTG 1594
Db 1199 CCACCTGTGTCAACCTCTGTAGTGTGTACAAACCAACAGCGGACCTTGCAGTAT 1258
QY 1595 CTGTCAAAAATATCAAGATGTGAGCTGGAAGCTTCGAGGGCACTGGAGTGTAC 1654
Db 1259 GTGGCGTGGATGTAAAGATGTGACTTCGAGAACTTCGGCATCCGTGGGTGTGTCT 1318
QY 1655 CTCAGATGCTGTCTCTTCATATACTATTTATTAACCTCTTATATGAAAAACATCA 1714
Db 1319 TCCAGGAGCGCAGCTTGTCTCTGTGATTTATTCAGAAAAACATCGCTACGGTCCAG 1378
QY 1715 GTGCTTCACTGAGGAAGTGTATGCAAGTGGCAAAATTAGTGAATTGATGATCAATTC 1774
Db 1379 GTGCCACCCGGAGAGATCATCGAAGTGGCTAAGAAAGCAAGCAATGATGATTCATTT 1438
QY 1775 TTCGAATGCCATGATATATACACCCAAAGTGGGAAAGAGACTCAAGCTTTCAGGAG 1834
Db 1439 CCGCTTCCCTGAAGATATGAAACCGTGTGGTGAACGCGGACTCAAACTTCTGTGTG 1498
QY 1835 GAGAAAAGCAAAAGATGCAATTTGCAAGAGCCATTTTGAAGACCCCGAGTCACTACT 1894
Db 1499 GCCAGAAAGACGCGCTCTCTGTGGACGCGGCAATGCTTAAGATGCCCATCTTCTGTTC 1558
QY 1895 ATGATGAAGTACTTCACTCGTTAGATTGATTAAGTGAAGACTATTCCTGTGCCATGA 1954
Db 1559 TCGATGAAGCCACTGTGCACTGATACCAAGTCTGAGCAGCAGTCCAAGCGGTTGG 1618
QY 1955 AGGATGTGTCAAAACAGAACTTCTATTTTCATTGCAACAGATTGTCAACAGTGTG 2014
Db 1619 AACAGCTGATGAAAAACGCACTTAATGATCGCCACCGCTGTCCACATCGCAG 1678
QY 2015 ATGAGATGAATATCATTTGTGATCAGGTTAGTAGCCGA 2057
Db 1679 GCGTCAATACCATGTGATCCATAAAAGGACGGGTTGAAGA 1721
```

Search completed: February 15, 2003, 05:51:32
Job time : 148 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 04:48:00 ; Search time 24 Seconds
(without alignments)
8105.166 Million cell updates/sec

Title: AF133659
Perfect score: 4254
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p model -DEV=xlh
-Q=/cgr2_1/USFTO_spool/AF133659/runat_10022003.155440.29513/app_query.fasta_1.2503
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum2 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcp -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=AF133659 @CGN 1.1.25 @runat_10022003.155440.29513 -NCPU=6 -ICPU=3
-NO_XLPLY -NO_WMAP -LARGESQDBY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WAPN TIMEOUT=30 -THREADS=1 -KAPOP=10 -XGAPOP=6 -FGAPEXT=7
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	ID	Description
1	3836	90.2	752	1	ABC7 HUMAN
2	3342	78.6	694	1	ABCT MOUSE
3	1538.5	36.2	693	1	YET1 SCHPO
4	1516.5	35.6	690	1	ATM1 YEAST
5	1239.5	22.9	842	1	ABCB HUMAN
6	973.5	15.9	830	1	HMT1 SCHPO
7	674.5	15.8	707	1	YFIC BACSU
8	673.5	15.8	707	1	HLV2 ECOLI
9	671.5	15.8	707	1	HLV2 ECOLI
10	664.5	15.6	715	1	AB10 ECOLI
11	660	15.5	708	1	HLV2 PASHA
12	654	15.4	586	1	Y4GM RHISN
13	651.5	15.3	707	1	HLV2 ACTAC
14	651	15.3	582	1	MSBA ECOLI
15	648	15.1	708	1	HLV2 PASSP
16	641.5	15.1	707	1	RT1B ACTPL
17	634	14.9	712	1	CYAB BORPE
18	632	14.9	587	1	MSBA_HAEIN

19	631.5	14.8	707	1	HLV2 PROVU
20	618.5	14.5	711	1	RT1B ACTPL
21	618	14.5	575	1	YWJA BACSU
22	613	14.4	738	1	AB10 HUMAN
23	613	14.4	1302	1	MDR4 DROME
24	611	14.4	614	1	YAS1_HAEIN
25	610	14.3	631	1	YC72 MYCTU
26	601.5	14.1	598	1	Y288 THEMEA
27	596.5	14.0	1276	1	MDR2 CRIGR
28	592.5	13.9	717	1	COMA STRPN
29	586.5	13.8	762	1	ABG3 MOUSE
30	586	13.8	762	1	ABG3 RAT
31	585	13.8	1276	1	MDR1 MOUSE
32	582.5	13.7	766	1	ABG9 HUMAN
33	582.5	13.7	1278	1	MDR2 RAT
34	579	13.6	1302	1	MDR5 DROME
35	577	13.6	1302	1	MDR5 RAT
36	572.5	13.5	1276	1	MDR2 MOUSE
37	572.5	13.5	1280	1	MDR1 HUMAN
38	572.5	13.5	1281	1	MDR3 CRIGR
39	571.5	13.4	735	1	ABG8 HUMAN
40	571	13.4	715	1	LCNC LACLA
41	570.5	13.4	589	1	Y015 MYCGE
42	570.5	13.4	722	1	MESD LEUME
43	569.5	13.4	607	1	HEPA ANASP
44	568.5	13.4	859	1	YD48 MYCTU
45	568.5	13.4	1279	1	MDR3_HUMAN

ALIGNMENTS

RESULT 1
ID ABC7 HUMAN STANDARD: PRT: 752 AA.
AC 075027; 075345; Q9UND1; Q9P0P1; Q9BRE1;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP-binding cassette, sub-family B, member 7, mitochondrial precursor
GN (ATP-binding cassette transporter 7) (ABC transporter 7 protein).
GN ABCB7 OR ABC7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98284536; PubMed=9621516;
RA Shimada Y., Okuno S., Kawai A., Shinomiya H., Saito O., Suzuki M.,
RA Omori Y., Nishino N., Kanemoto N., Fujiwara T., Horie M.,
RA Takahashi E.,
RT "Cloning and chromosomal mapping of a novel ABC transporter gene
RT (hABC7), a candidate for X-linked sideroblastic anemia with
RT spinocerebellar ataxia.";
RL J. Hum. Genet. 43:115-122 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=98318631; PubMed=9653160;
RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
RA He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
RA Wang Y.-X., Chen S.-J., Chen Z.-J.
RT "Identification of genes expressed in human CD34(+) hematopoietic
RT stem/progenitor cells by expressed sequence tags and efficient full-
RT length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180 (1998).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT ASAT MET-400.
RX MEDLINE=99214014; PubMed=10196363;
RA Allkmeers R., Raaskind W.H., Hutchinson A., Schneck N.D., Dean M.,
RA Koeller D.M.;
RT "Mutation of a putative mitochondrial iron transporter gene (ABCT7) in

RT X-linked sideroblastic anemia and ataxia (XLSA/A).";
 RL Hum. Mol. Genet. 8:743-749(1999).
 RN [4]
 RN SEQUENCE FROM N.A., AND VARIANT ASAT LYS-433.
 RP MEDLINE=20504074; PubMed=11050011;
 RX Bekri S., Kispal G., Lange H., Fitzsimons E., Tolmie J., Lill R.,
 RA Bishop D.F.;
 RT "Human ABC7 transporter: gene structure and mutation causing X-linked
 RT sideroblastic anemia with ataxia with disruption of cytosolic
 RT iron-sulfur protein maturation.";
 RL Blood 96:3256-3264(2000).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RC Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE OF 5-752 FROM N.A.
 RP MEDLINE=99098366; PubMed=9883897;
 RX Casare P., Lill R., Kispal G.;
 RT "Identification of a human mitochondrial ABC transporter, the
 RT functional orthologue of yeast Atm1p.";
 RL FEBS Lett. 441:266-270(1998).
 CC -1- FUNCTION: COULD BE INVOLVED IN THE TRANSPORT OF HEME FROM THE
 CC MITOCHONDRIA TO THE CYTOSOL. PLAYS A CENTRAL ROLE IN THE
 CC MATURATION OF CYTOSOLIC IRON-SULFUR (FE/S) CLUSTER-CONTAINING
 CC PROTEINS.
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (Potential).
 CC -1- DISEASE: DEFECTS IN ABCB7 ARE THE CAUSE OF X-LINKED SIDEROBLASTIC
 CC ANEMIA WITH ATAXIA (ASAT). ASAT IS A RECESSIVE DISORDER
 CC CHARACTERIZED BY AN INFANTILE TO EARLY CHILDHOOD ONSET OF
 CC NONPROGRESSIVE CEREBELLAR ATAXIA AND MILD ANEMIA WITH HYPOCHROMIA
 CC AND MICROCYTOSIS.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL, AB005289; BAA28861.1; -;
 DR EMBL, AF038950; AAC39865.1; -;
 DR EMBL, AF133659; AAD33046.1; -;
 DR EMBL, AF241887; AAK20173.1; -;
 DR EMBL, AF241872; AAK20173.1; JOINED.
 DR EMBL, AF241873; AAK20173.1; JOINED.
 DR EMBL, AF241874; AAK20173.1; JOINED.
 DR EMBL, AF241875; AAK20173.1; JOINED.
 DR EMBL, AF241876; AAK20173.1; JOINED.
 DR EMBL, AF241877; AAK20173.1; JOINED.
 DR EMBL, AF241878; AAK20173.1; JOINED.
 DR EMBL, AF241879; AAK20173.1; JOINED.
 DR EMBL, AF241880; AAK20173.1; JOINED.
 DR EMBL, AF241881; AAK20173.1; JOINED.
 DR EMBL, AF241883; AAK20173.1; JOINED.
 DR EMBL, AF241883; AAK20173.1; JOINED.
 DR EMBL, AF241884; AAK20173.1; JOINED.
 DR EMBL, AF241885; AAK20173.1; JOINED.
 DR EMBL, AF241886; AAK20173.1; JOINED.
 DR EMBL, BC006323; AAH06323.1; -;
 DR EMBL, AF078777; AAD47141.1; -;
 DR Genew; HGNC:48; ABCB7.
 DR MIM; 300135; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR001140; ABCtranspLtrTM.
 DR Pfam; PF00005; ABC_tran; 2.

DR	Pfam: PF00664; ABC membrane; 2.	
DR	ProDom: PD00606; ABC transport; 1.	
DR	SMART: SM00382; AAA; 1.	
DR	PROSITE, PS00211; ABC_TRANSPORTER; 1.	
KW	ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane.	
KW	Transit peptide; Disease mutation.	
FT	TRANSIT	1
FT	CHAIN	? 752
FT		
FT	TRANSMEM	260 280
FT	TRANSMEM	291 311
FT	TRANSMEM	383 403
FT	NP_BIND	505 512
FT	VARIANT	400 400
FT		
FT	VARIANT	433 433
FT		
FT	CONFLICT	56 56
FT	CONFLICT	141 141
FT	CONFLICT	258 258
FT	CONFLICT	271 276
FT	CONFLICT	281 281
FT	CONFLICT	290 290
FT	CONFLICT	293 297
FT	CONFLICT	315 315
FT	CONFLICT	320 324
FT	CONFLICT	346 346
FT	CONFLICT	542 542
Q	SEQUENCE	752 AA; 82641 MW; B1FEA57ABD24FB90 CRC64;

```

AF133659 (1-2345) x ABC7_HUMAN (1-752)

Alignment Scores:
Pred. No.:      7,96e-855      Length:      752
Score:          3836.00         Matches:      752
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%   Mismatches:  0
Query Match:    90.17%         Indels:      0
DB:             1              Gaps:        0

```

[illegible]

QY 481 TTTAATATGCTGATGACAGCTCAACAGATGTCGGGAAACATGCTGAACCTGAT 540
 DB 161 PheylsYrAlaValAspSerLeuasnInleuSerGlyAsnMetLeuAsnLeuSerAsp 180
 QY 541 GCACCAATATGATGACACCATGGCAACAGATTCATGTTGGCTATGATGATCAAGA 600
 DB 181 AlaProAsnThrValAlaThrMetAlaThrAlaValLeuIleGlyTYrGlyValSerArg 200
 QY 601 GCTGAGCTGCTTTTAAACGAAGTTGAAATGCAATGATTTGGCAAGTGAACCCAGAT 660
 DB 201 AlaGlyAlaAlaPhePheAsnGluValArgAsnAlaValPheGlyLysValAlaGlnAsn 220
 QY 661 TCATCCGAAATAATGCAAAATGCTTTCTCCATCTTCACACCTGATCTGGGTTT 720
 DB 221 SerIleAspArgIleAlaLysAsnValPheLeuHisLeuHisAsnLeuAspLeuGlyPhe 240
 QY 721 CACCTGAGCAGACAGACGGAGCTTATCTAAGCTATGACAGAGAAACAGGGATGATC 780
 DB 241 HisLeuSerArgGlnThrGlyAlaLeuSerLysAlaIleAspArgGlyThrArgGlyIle 260
 QY 781 AGTTTGTCTGAGTGTCTTGGATTTAATCTTCTCCATCATGTTGAAGTATGCTT 840
 DB 261 SerPheValLeuSerAlaLeuValPheAsnLeuLeuProIleMetPheGluValMetLeu 280
 QY 841 GTCAGTGGTGTGTTGATTAACAATGCGGTGCCAGTTGCTTGGTACCCTTGGAACA 900
 DB 281 ValSerGlyValLeuTYrTYrLysCysGlyAlaGlnPheAlaLeuValThrLeuGlyThr 300
 QY 901 CTGATGATCATGACAGATTCACAGATTCAGAGTTCACAGGTGGAGACATGATTAAGATA 960
 DB 301 LeuGlyThrTYrThrAlaPheThrValAlaValThrArgThrArgThrArgPheArgIle 320
 QY 961 GAAATGAACAAAGCAGATTAATGATGACAGTATGCTGCTATAGACTCACTGCTGAATTA 1020
 DB 321 GluMetAsnLysAlaAspAsnAspAlaGlyAsnAlaAlaIleAspSerLeuLeuAsnTYr 340
 QY 1021 GAAATGATGATTTTAAATGAAGAATATGAGACACAGACATATGATGATTTTGG 1080
 DB 341 GluThrValLysTYrPheAsnAsnGluArgTYrGluAlaGlnArgTYrAspGlyPheLeu 360
 QY 1081 AAGAGTATGAGACTGCTTCACTGAAAGTCTCTACTGCTGATGCTGAATCTTGGT 1140
 DB 361 LysThrTYrGluThrAlaSerLeuLysSerThrSerThrLeuAlaMetLeuAsnPheGly 380
 QY 1141 CAAGTCTATTTTCAAGTGTGCTTTAAACAGCTAATGATGCTGCGCAGTCAAGGAATT 1200
 DB 381 GlnSerAlaIlePheSerValGlyLeuThrAlaIleMetValLeuAlaSerGlnGlyIle 400
 QY 1201 GTGGCAGGATCCTTACTGTTGAGATCTAGTAATGATGTAATGGAATGCTTTTCACTT 1260
 DB 401 ValAlaGlyThrLeuThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeu 420
 QY 1261 TCATACCCCGAATCTTTCGGGAATCTGATATGAGAGATGACACAGCACTCATGAT 1320
 DB 421 SerLeuProLeuAsnPheLeuGlyThrValTYrGluThrArgGlnAlaLeuIleAsp 440
 QY 1321 ATGAACACCTTGTACTCTACTCAAGATGACACCAACAAATTAAGCAAAAGTATGCA 1380
 DB 441 MetAsnThrLeuPheThrLeuLeuLysValAspThrGlnIleLysAspLysValMetAla 460
 QY 1381 TCTGCCCTTCAATCACACACAGACAGCTCCGTGCTTGTGAATATGTCATTTGAA 1440
 DB 461 SerProLeuGlnIleThrProGlnThrAlaThrValAlaPheAspAsnValHisPheGlu 480
 QY 1441 TACATTGAGGGCCAGAAAGCTTATGTAAGTAATCTTTGAAGTCCCTGCAGAGAAAGAAA 1500
 DB 481 TYrIleGlnGlyGlnLysValLeuSerGlyLysSerPheGlnValProAlaGlyLysLys 500
 QY 1501 GTGGCCATTTAGAGAGTAGTGGGTCAAGGAAAAACAAATAGAGAGGCTATTATTTCCG 1560
 DB 501 ValAlaIleValGlyLysSerGlySerGlyLysSerThrIleValAlaGluLeuPheArg 520
 QY 1561 TTCTATAGCCTCAAAAAGGAGTATTAATCTTGTGCTGCAAAAATATACAAAGATGTGAGC 1620

DB 521 PheTYrGluProGlnLysGlySerIleTYrLeuAlaGlyGlnAsnIleGlnAspValSer 540
 QY 1621 CTGAAAGCCTTCGAGAGGAGGAGGTGGAGTGTACTCCAGAGATCTGCTTCATTAAT 1680
 DB 541 LeuGlnSerLeuArgArgAlaValGlyValValProGlnAspAlaValLeuPheHisAsn 560
 QY 1681 ACTATTTATTACAACCTCTTATATGGAACATGACGTCTTCACTGAGGAAGTATGCA 1740
 DB 561 ThrIleTYrTyrsnLeuLeuTYrGlyAsnIleSerAlaSerProGlnGluValTYrAla 580
 QY 1741 GTGGCAAAATTTAGCTGACTTCATGATGCAATTTCTGAAATGCCACATGATATGACAC 1800
 DB 581 ValAlaLysLeuAlaGlyLeuHisAspAlaIleLeuArgMetProHisGlyTYrAspThr 600
 QY 1801 CAAGTGGGGAACGAGACCTCAAGCTTTCAGAGAGAGAAAGCAAGAGATGACATTTGCA 1860
 DB 601 GlnValGlyGluArgGlyLeuLysLeuSerGlyGlyGlnLysGlnArgValAlaAlaIleAla 620
 QY 1861 AGAGCATTTTGAAGAACCCCGCATATCTATGATGAGACCTATCTCATGCTTATGAT 1920
 DB 621 ArgAlaIleLeuLysAspProProValIleLeuTYrAspGlnAlaThrSerSerLeuAsp 640
 QY 1921 TCGATTACTGAAGACATATTTCTTGGTCCATGAGATGTGTCAAACACAGAACTTCT 1980
 DB 641 SerIleThrGlnGluThrIleLeuGlyAlaMetLysAspValValLysHisArgThrSer 660
 QY 1981 ATTTTCATTCGACACAGATTTGTCACAGTGGTTGAGCAGATGAATTCATTTGCTGGAT 2040
 DB 661 IlePheIleAlaHisArgLeuSerThrValValAspAlaAspIleIleValLeuAsp 680
 QY 2041 CAGGTGAGGTAGCCGAAAGTGTGATCCCAACATGTTGCTTGTCTAACCCTCATGATATC 2100
 DB 681 GlnGlyLysValAlaGluArgGlyThrHisGlyLeuLeuAlaAsnProHisSerIle 700
 QY 2101 TATTCAAAATGTGCAATACACAGACAGCCGTGTGCAAGACATGATTAACCCCAATG 2160
 DB 701 TYrSerGluMetIleThrHisThrGlnSerSerArgValGlnAsnHisAspAsnProLys 720
 QY 2161 GAAGCAAGAGAAAGAAATTAATCAAGAGAGAGAAAGAAAGAAATTCACAGAAAGAAAT 2220
 DB 721 GluAlaLysLysGlnAsnIleSerLysGlnGluGlnArgLysLysLeuGlnGluIle 740
 QY 2221 GTCAATAGTGTGAAGAGCTGTGGAACCTGTCGTC 2256
 DB 741 ValAsnSerValLysGlyCysGlyAsnCysSerCys 752

RESULT 2
 ABC7 MOUSE
 ID ABC7 MOUSE STANDARD; PRT; 694 AA.
 AC 061102;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE ATP-binding cassette, sub-family B, member 7, mitochondrial (ATP-binding cassette transporter 7) (ABC transporter 7 protein) (Fragment).
 GN ABC7 OR ABC7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1 TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=DBA/2;
 RX MEDLINE=97288528; PubMed=9143506;
 RA Savary S., Allikmets R., Denizot F., Luciani M.-F., Mattei M.-G., Dean M., Chimini G.,
 "Isolation and chromosomal mapping of a novel ATP-binding cassette transporter conserved in mouse and human."
 RL Genomics 41:275-278(1997).
 CC -!- FUNCTION: COULD BE INVOLVED IN THE TRANSPORT OF HEME FROM THE MITOCHONDRIA TO THE CYTOSOL. PLAYS A CENTRAL ROLE IN THE

CC	MATURATION OF CYTOSOLIC IRON-SULFUR (FE/S) CLUSTER-CONTAINING PROTEINS (BY SIMILARITY).
CC	-1- SUBUNIT: HOMODIMER OR HETERODIMER (POTENTIAL).
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (Potential).
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-slb.ch/announce/ or send an email to license@isb-slb.ch).
CC	-----
DR	EMBL; U43892; AAC53152.1; -.
DR	MCD; MG1:109533; Abcb7.
DR	InterPro; IPR003593; AAA_Arpase.
DR	InterPro; IPR003439; ABC_Transporter.
DR	InterPro; IPR001140; ABCTransportrTM.
DR	Pfam; PF00005; ABC_Tran_1.
DR	Pfam; PF00664; ABC_membrane_1.
DR	PfDomam; PD000006; ABC_transpoitr; 1.
DR	SMART; SMO0382; AAA; 1.
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW	ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane.
FT	NON TER 1 1
FT	TRANSMEM 202 222 POTENTIAL..
FT	TRANSEM 233 253 POTENTIAL.
FT	TRANSEM 326 346 POTENTIAL.
NP_BIND	447 454 ATP (POTENTIAL).
QO	SEQUENCE 694 AA; 76417 MW; A7AE89EAE9AA581D CRC64;

[illegible]

Oy	1717	GCTTCACTGAGGAAGTGATGCAGTGGCAAAATTAGTCGACCTCATGATGCAATCTT	1776
Dd	544	AAlaSerAsnAspIuLeValIGuaIaAbAlaYslySAlaYsIIeHisAspIIeIleGu	563
Oy	1777	CGAATGCCACATGATATATACACCACAAGTAGGGGAACAGACACTCACAACCTTTACAGAGA	1836
Dd	564	SerPheProGluGlyYrGlnThrIysValIGlyIuArgGlyLeuMetIISeSerGlyCy	583
Oy	1837	GAAGAAGCAAGATAGCAATTTGGCAAGACCACTTTTGAAGAGACCCCCAGTCATCTTAT	1896
Dd	584	GIuLySGInArgLeuAlaValSerArgLeuLeuLysAsnProGluIleLeuPheHe	603
Oy	1897	GATGAACCTCTTCATCCGTGTAAGTTACTGTAAGACACTATTCTTTGGTCCATGAAG	1956
Dd	604	AspGIuAlaThrSerAlaLeuAspThrAsnThrGluAlaLeuAlaLeuAlaAsnIleSn	623
Oy	1957	GATGTGGTCAA-----CACGAACCTTATTTTCATGGCACACAGATTGTCCAACAGTG	2010
Dd	624	AspLeuIleYsGlySerHisIyStrHisSerValPheIleIaHisArgLeuArgThrIle	643
Oy	2011	GTTTGATGCAGATGAATTCATTGTCTTGATCAGGGGTAAAGTAGCCGAAGTGGTACCAC	2070
Dd	644	LysAspCysAspIleIlePheValLeuGluYsGlyArgValIaIGluGlnGlySerHis	663
Oy	2071	CATGCTTTGCTGCTAACCTCATAGTATCTTTCAGAAATGTGGCATACACAGACAGC	2130
Dd	664	GIuIuIleuIleuIleuIleuIleu--AsnSerValYrTrpSerMetITrpHisSerGlnGlySer	682

```

RESULT 4
ID      ATML1_YEAST
AC      P40416
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Transporter ATML1, mitochondrial precursor.
GN      ATML1 OR MDY OR YMR301C OR YM9952.03C.
OS      Saccharomyces cerevisiae (Baker's Yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxID=4932;
LN      [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=JK9-3D ALPHA.
RC      MEDLINE=95129546; PubMed=7828591;
RX      Leighton J., Schatz G.;
RA      "An ABC transporter in the mitochondrial inner membrane is required
RT      for normal growth of yeast."
RL      EMBO J. 14:188-195(1995).
RN      [2]
RN      SEQUENCE FROM N.A.
RP      MEDLINE=98089018; PubMed=9428742;
RX      Kispal G., Csere P., Guiard B., Lill R.;
RA      "The ABC transporter Atm1p is required for mitochondrial iron
RT      homeostasis."
RL      FEBS Lett. 418:346-350(1997).
RN      [3]
RN      SEQUENCE FROM N.A.
RP      STRAIN=S288c / AB972;
RC      Connor R., Chuber C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA      Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: PROBABLE TRANSPORTER FOR A YET UNKNOWN SUBSTRATE.
CC      REQUIRED FOR MITOCHONDRIAL IRON HOMEOSTASIS.
CC      -1- SUBUNIT: HOMODIMER OR HETERODIMER (POTENTIAL).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC      inner membrane.
CC      -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL, on station -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```


Qy 223 GGAAGCTTTTGAATGCTGCAAGGCTCTCCAGGTATGCCATGTGATGAAGAAAGACACA 282
Db 184 ValGlnPhe-----SerLeuTrpValLeu----- 191
Qy 283 TGTGTGGCATGTCATGCAGAGGAGGACTCCACACA----- 318
Db 192 -----ArgTyrValValSerGlyGlyLeuPheValLeuGlyLeuTrpAlaProGlyLeu 209
Qy 319 -----GACCCAAAAGAGGGTTAAAGATGTGATGACTCGGAAA 357
Db 210 ArgProGlnSerTyrThrLeuGlnValHisGluGluAspGlnAspValGluAspSerGln 229
Qy 358 ATC-----ATTAAGCAATG 372
Db 230 ValArgSerAlaAlaGlnGlnSerThrTrpArgAspPheGlyArgLysLeuArgLeuLeu 249
Qy 373 CTTTCTATATGTCGCCCCAAAGACAGCCAGATCTACAGCTAAGTTGCACTTTCGCTG 432
Db 250 SerGlyTyrLeuTrpProArgGlySerProAlaLeuGlnLeuValValLeuIleCysLeu 269
Qy 433 GGAATTTTGGTGGTGCAGAAAGCCATGATATGTGGTCCCTTCATGTTTAAATATGCT 492
Db 270 GlyLeuMetGlyLeuGluArgAlaLeuHisValLeuValProIlePheTyrArg----- 287
Qy 493 GTAGACAGCCTCAACACAGATGTCGGAAACATGCTGAACCTG-----AGTATGCACCA 546
Db 288 -----AsnIleValAsnLeuLeuThrGluLysAlaPro 298
Qy 547 ---AATACAGTGCACACCATGCGACACACA-----GTTTCGATTTGGCTAT 588
Db 299 TrpSerSerLeuAlaTrpThrValThrSerTyrValPheLeuLysPheLeuGlnGly 318
Qy 589 GGTGTATCAAGAGCTGAGCTGCTTTTATTAACGAATGTGCAAGTGCAGTATTTGGCAG 648
Db 319 GlyThrGlySerThrGly-----PheValSerAsnLeuHisGlyThrPheLeuTrpIleArg 336
Qy 649 GTAGCCCAAGATTCAATCCGAGAAATAGCCAAAATGTCTTTCATCTTCACCAACCTG 708
Db 337 ValGlnGlnPheThrSerArgArgValGluLeuLeuIlePheSerHisLeuHisGluLeu 356
Qy 709 GATTTGGGTTTTCACCTCGAGACACAGACGGAGCTTATGTTAAAGCTATTTAGACAGGA 768
Db 357 SerLeuArgTrpHisLeuGlyArgArgThrGlyValLeuAlaIleAlaAspArgGly 376
Qy 769 ACAAGGGGTATCAGTTTGTCTGAGTGTGTTGGTATTTATATCTTCCCATGATGTTT 828
Db 377 ThrSerSerValThrGlyLeuLeuSerTyrLeuValPheAsnValIleProThrLeuAla 396
Qy 829 GAAATGATGCTGTGTCAGTGTGTTTGTATTAACAATGC-----GGTGGCCAGTTTGT 882
Db 397 AspIleIleIle-----GlyIleIleTyrPheSerMetPhePheAsnAlaTrpPheGly 414
Qy 883 TTGGTAACCTTGGACACACTTGGTACATPACACAGCACTTACAGTTGACGTACACGGTGG 942
Db 415 LeuIleValPheLeuCysMetSerLeuTyrLeuThrIleValIleValThrGluTrp 434
Qy 943 AGAAGTGAATTTGAATGAAGAATGAACAGACAGATATGATGAGGTAATGCTGCTATA 1002
Db 435 ArgThrLysPheAspArgAlaMetAsnThrGlnGluAsnAlaThrArgAlaArgAlaVal 454
Qy 1003 GACTCACTGCTGAATTTGAAGAACTGTGAAGTATTTATATGAAGAAATGAAGACAGAC 1062
Db 455 AspSerLeuLeuAsnPheGluThrValLysTyrThrAsnAlaGlnSerTyrGluValGlu 474
Qy 1063 AGATATGATGATTTTGAAGACGTATGACAGCTGCTTCATTGAAGAAAGTACTTACTCTG 1122
Db 475 ArgTyrArgGluAlaIleIleTyrGlnGlyLeuGluTrpLysSerSerAlaSerLeu 494
Qy 1123 GCTATGCTGAACCTTGTGTCGAAGGTATTTTTCAGTGTGCTTAAACAGCTATATGCTG 1182
Db 495 ValLeuLeuAsnGlnThrGlnAsnLeuValIleGlyLeuGlyLeuLeuAlaGlySerLeu 514

Qy 1183 CTGGCCAGTCAAGGAATTTGGCGAGGTACCTTACTTGTGGAGATCTAGATATGCTGAAT 1242
Db 515 LeuCysAlaTyrPheValThrGluGlnLysLeuGlnValGlyLysPyrValLeuPheGly 534
Qy 1243 GGACTGCTTTTTCAGCTTTTCAATTAACCCCTGAACCTTCTGGAACTGATATAGAGACT 1302
Db 535 ThrTyrIleIleGlnLeuTyrMetProLeuAsnTrpPheGlyThrTyrTyrArgMetIle 554
Qy 1303 AGACAAGCACTCATGATATGAACACCTGTTTACTTACTTACAGGTAGACACCCAAAT 1362
Db 555 GlnThrAsnHelleAspMetGluAsnMetPheAspLeuLeuLysGluGluVal 574
Qy 1363 AAAGACAAAGTATGATGATGATCCCTTCAGATCACACACAGACAGCTACCGTGGCCTT 1422
Db 575 LysAspLeuProGlyAlaGlyProLeuArgPhe-----GlnLysGlyArgIleGluPhe 592
Qy 1423 GATTAATGCTATTTGAATCATTCATTGAGGGCGACAAATCCTTACTGTAATATCCTTTGAA 1482
Db 593 GluAsnValHisPheSerTyrAlaAspGlyArgGluThrLeuGlnAspValSerPheThr 612
Qy 1483 GTCCCTGCAGAAAGAAAGTGGCCATTTGAGAGTGTGAGGTGCAGGGAAAGACACATA 1542
Db 613 ValMetProGlyGlnThrLeuAlaLeuValGlyProSerGlyAlaGlyLysSerThrIle 632
Qy 1543 GTGAGCTATTTATTTGCTTCTATGAGCTCAAAAGGGTACATTATCTTGCTGTCAA 1602
Db 633 LeuArgLeuLeuPheArgPheTyrAspIleSerSerGlyCysIleArgIleAspGlyGln 652
Qy 1603 AATATACAAATGTGAGCTCGGAAAGCTTCGAGGCACTGGAGTGTGATCTCAGAT 1662
Db 653 AspIleSerGlnValThrGlnAlaSerLeuArgSerHisIleGlyValValProGlnAsp 672
Qy 1663 GCTGCTCTTTCATTAATCTATTTATTAACACCTCTTATGGAACACAGTCTTCA 1722
Db 673 ThrValLeuPheAsnAspThrIleAlaAspAsnIleArgTyrGlyArgValThrAlaGly 692
Qy 1723 CTTGAGGAATGTATGACGTGGCAAAATTAAGCTGACCTTATGATGATGATG 1782
Db 693 AsnAspGluValGluAlaAlaIleAlaGlnAlaAlaGlyIleHisAspAlaIleMetAlaPhe 712
Qy 1783 CCACATGATATGACACCCCAAGTGGGAAACGAGACTCAAGCTTTTACAGAGGAAAG 1842
Db 713 ProGlnGlyTyrArgThrGlnValGlyGluArgGlyLeuLysLeuSerGlyGlyGlyLys 732
Qy 1843 CAAGAAGTACCAATTTGCAAGACCTTTTGAAGACCCCCAGCTACTCTATGATGAA 1902
Db 733 GlnArgValAlaIleAlaAspGlyThrIleLeuLysAlaProGlyIleIleLeuAspGlu 752
Qy 1903 GCTACTTCATCGTTAGATTCGATTACTGAAGACATATCTTGTCGCATGAAGATGTG 1962
Db 753 AlaThrSerAlaLeuAspThrSerAsnGluArgAlaIleGlnAlaSerLeuAlaLysVal 772
Qy 1963 GTCAACACAGAACTTCTATTTTTCATTTGCACACAGATTGTCAACAGTGTGATGCAAT 2022
Db 773 CysAlaAsnArgThrThrIleValValAlaHisArgLeuSerThrValValaAsnAlaAsp 792
Qy 2023 GAAATGATGCTTGTGATGACGGTAAAGTACCGCAAGCGTGAACCACTATGTTGCTT 2082
Db 793 GlnIleLeuValIleLysAspGlyCysIleValGluArgGlyArgHisGluAlaLeuLeu 812
Qy 2083 GCTAACCTCATGATCTATTCATGAGAATGTGCATACACAGACAGCCGTGTGCAGAAC 2142
Db 813 SerArg---GlyGlyValTyrAlaAspMetTrpGlnLeuGlnGlnGlyGlnGluThr 831
Qy 2143 CATTGATACCCCAA 2157
Db 832 SerGluAspThrLys 836
RESULT 6
HMT1_SCHPO
ID HMT1_SCHPO STANDARD: PRT; 830 AA.
AC 002592; 013675; 090QW7; 090S13;
01-JUL-1993 (Rel. 26, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heavy metal tolerance protein precursor.
 GN HMT1 OR SPCC74.08C OR SPCC37.09C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NC NCB1 TaxID=4896;
 RX MEDLINE=972;
 RP STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymopre B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filtz C., Holzer E., Moesli D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Mambutt R., Punelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hure S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Sherkovskii G.V., Ussery D., Barrett B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: INVOLVED IN METAL TOLERANCE. PROBABLY INVOLVED IN THE
 CC TRANSPORT OF METAL-BOUND PHYTOCHELATINS. COMPARTMENTALIZES
 CC CADMIUM WITHIN VACUOLES, THEREBY PROTECTING CELLS FROM CADMIUM
 CC TOXICITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; Z14055; CAJ78419.1; -;
 DR EMBL; AL031543; CAJ20838.1; -;
 DR EMBL; AL031546; CAJ20865.1; -;
 DR PIR; S25198; S25198.
 DR HSP; P13569; INBD.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR01140; ABCtransporter.
 DR Pfam; PF00064; ABC_tran; 1.
 DR Pfam; PF00064; ABC_membrane; 1.

DR ProDom; PD00006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Transport; Transmembrane; Signal; ATP-binding; Glycoprotein;
 KM Cadmium resistance.
 FT SIGNAL 1 27
 FT CHAIN 28 830 HEAVY METAL TOLERANCE PROTEIN.
 FT TRANSMEM 10 31 POTENTIAL.
 FT TRANSMEM 52 71 POTENTIAL.
 FT TRANSMEM 88 108 POTENTIAL.
 FT TRANSMEM 126 146 POTENTIAL.
 FT TRANSMEM 403 423 POTENTIAL.
 FT TRANSMEM 490 511 POTENTIAL.
 FT NP BIND 617 624 ATP (POTENTIAL).
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 79 79 R -> A (IN REF. 1).
 FT CONFLICT 439 439 S -> T (IN REF. 1).
 FT CONFLICT 812 812 A -> R (IN REF. 1).
 SQ SEQUENCE 830 AA; 93993 MW; 909FBD10D51F50A9 CRC64;
 Alignment Scores:
 Pred. No.: 4,2e-59 length: 830
 Score: 973.50 Matches: 220
 Percent Similarity: 56.97% Conservative: 115
 Best Local Similarity: 37.41% Mismatches: 236
 Query Match: 22.88% Indels: 17
 DB: 1 Gaps: 7
 AF133659 (1-2345) x HMT1_SCHPO (1-830)
 QY 370 ATGCTTCTTATGATGAGCC---AAAGAC---AGCCAGATCTACGAGATTGCC 423
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 249 LEUENPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRY 268
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 424 ATTTGCGTGGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 483
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 269 ILEVALLLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUT 286
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 484 AATATGCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 287 GLYVALLLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUTPOTRYLEUT 304
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 544 CCAATACAGTTGACACCATGACACCATGACACCATGACACCATGACACCATGACAC 603
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 305 -----Valleuubhevalleuubhevalleuubhevalleuubhevalleuubhe 316
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 604 GGAGCTGCTTTTAAAGAGATTGAGAGATTGAGAGATTGAGAGATTGAGAGATTGAG 663
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 317 AsmetgVallleuubhevalleuubhevalleuubhevalleuubhevalleuubhe 336
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 664 ATCCGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 723
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 337 TYRIGALALLEUUBHEVALLEUUBHEVALLEUUBHEVALLEUUBHEVALLEUUBHE 356
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 724 CTGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 783
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 337 LEUENLPSAGLALAGLGLVALLLEUTHRALLEUTHRALLEUTHRALLEUTHRAL 375
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 784 TTTGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 843
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 376 ThrphealagluginValleuubhevalleuubhevalleuubhevalleuubhe 395
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 844 AGTGCTTTTGTATTAATATGAGAGATTGAGAGATTGAGAGATTGAGAGATTGAGAG 903
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 396 MetValTYRphetheleuubhevalleuubhevalleuubhevalleuubheval 415
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 904 GTATCATACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 963
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 416 LeucTYRtyrTYRvalThvalThvalThvalThvalThvalThvalThvalThval 435
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
 QY 964 ATGAACAAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1023
 DB ::::: ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||

Db 436 MetValIenSerTPdAgIuSerTYrAlaValAlnIenAspAlaIleMetAsnPhcIu 455
 QY 1024 ACTGTGAATATTTAAATGAAGAATATGACAGACAGATATGATGATTTTGAAG 1083
 Db 456 ThrValIysAsnPhcAspAlaAspAspPhcIuIenGluArgIyRglYhIAlaValAsp 475
 QY 1084 ACGTATGAGACTGCTTATTCGAAAAGTACTTACTCTGCTGCTTACCTGAACTTGTCAA 1143
 Db 476 IleTYrLeuYsgIuArgIyValIeuPhcSerLeuIenAsnPhcAlaValAln 495
 QY 1144 AGTGTCTATTTTACGTGCGTTTAAAGCTTATGCGTCCGACAGCAGGAATGTG 1203
 Db 496 GlyGlyIlePheIenrPheSerLeuAlaIleAlaCySeuIeuSerAlaTYrArgValThr 515
 QY 1204 GCAAGTACCTTACTGTGAGATCTAGTATGAGATGAGTGAAGTCTTTTCACTTCA 1263
 Db 516 PheIlyPheAsnThrValIglYAspPheValIleIeuIeuThrTYrMetIleGlnIeuGln 535
 QY 1264 TTACCTCGAACTTTCTGGGAACTGTATATGAGAGACTAGACAGACCTCATGATATG 1323
 Db 536 GlnProIeuAsnPhcPheGlyThrLeuTYrArgSerLeuGlnAsnSerIleIleAspThr 555
 QY 1324 AACACCTTTTACTCTACTCAAGTATGACACCCAAATTAAAGCAAGATGATGATCT 1383
 Db 556 GluArgIeuIeuGluIlePheIgluIyPheProThrValValGluIyProAsnAlaPro 575
 QY 1384 CCCCTTCAGATCACACACAGACAGACTACCGCTTGTATATGATTTTGAATAC 1443
 Db 576 AspIeuYsValThr-----GlnGlyIyValIlePheSerIhIsvAlSerPheAlaTYr 593
 QY 1444 ATTGAGGCGCAAGAAATCTTATGGAATATCTTGAAGTCCCTGACAGGAAGAAAGT 1503
 Db 594 AspProArgIyProValIeuSerAspIleAsnPhcValAlaGlnProGlyIyValIle 613
 QY 1504 GCCATTGTGAGATGATGGTTCAGGAAAAGCATTAGAGAGCTATTTTCGCTTC 1563
 Db 614 AlaIeuValIglYIuSerGlyIyGlyIySerThrIleMetArgIleLeuIeuAlaArgPhe 633
 QY 1564 TATGAGCTCAAAAGGTAGATTCCTGCTGCTGCTCAAAATATACAGATGAGGCTG 1623
 Db 634 PheAspValIenSerGlySerIleThrIleAspAspIleAsnPhcValIleValThrIeu 653
 QY 1624 GAAAGCTTGGAGGAGGAGGAGGAGTGTACTGAGATGCTGCTTCATTAATATCT 1683
 Db 654 SerSerIeuArgSerIleGlyValIValProGlnAspSerThrIeuPhcAsnAspThr 673
 QY 1684 ATTATTTACAACTTATATGAAACATCAGTCTTCACTGAGGAGTATGCGATG 1743
 Db 674 IleIeuTYrAsnIleIyTYrAlaIyProSerAlaThrAsnIuIuIleTYrAlaAla 693
 QY 1744 GCAAAATTTAGCTGACTTATGATGATGAAATCTTGAATGCGACATGATGACACCCAA 1803
 Db 694 AlaIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 713
 QY 1804 GTAGGGAGACGAGACTCAAGCTTTTCAAGAGAGAAAAGCAAGATGAGATGACAGA 1863
 Db 714 ValGlyIuIuArgIyIeuIyIeuSerGlyIyGlyIyIuIyIuIyValAlaValAlaArg 733
 QY 1864 GCCATTTTGAAGACCCCACTATCTATGATGATGAGTACTTATGCTGATGATTCG 1923
 Db 734 AlaIleIeuYsAspProSerIleIleIleIeuAspIuIaIaThrSerAlaIeuAspThr 753
 QY 1924 ATTACTGAAGACTATTTCTGTGCGATGAGATGAGTGTGCAACACAGAACTTAT 1983
 Db 754 AsnThrGluArgIuIleGlnAlaIaIeuAsnArgIeuAlaSerGlyArgThrAlaIle 773
 QY 1984 TTCAATTCACACAGATTTGTCAACAGTGTGATGACAGATGAATATGCTTGGATCAG 2043
 Db 774 ValIleIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 793
 QY 2044 GGTAAAGTACCCCAAGCTGTGATCCACATGCTTGTCTGTAACCTCATATGATCTAT 2103
 Db 794 GlyArgIleValIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 813

QY 2104 TCAGAAATGTGCATACACAGAC 2127
 Db 814 IySeIySeIySeIyPheGlnIuAla 821
 RESULT 7
 YFIC_BACSU
 ID YFIC_BACSU STANDARD; PRT; 604 AA.
 AC P54719;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein yf1c.
 GN yf1c.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_Taxid:1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96262713; PubMed=8704981;
 RA Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;
 RT "Determination of a 12 kb nucleotide sequence around the 76 degrees
 RL Microbiology 142:1417-1421 (1996).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Besseires P., Bolotin A., Borchert S.,
 RA Bortiss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch J., Caldwell B., Camargo N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings V.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entlin K.D., Errington J., Fabre C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Guim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Hatech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Jorj B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber V., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Putrelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan B., Schleich S., Schroeder R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seir S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Toseco V., Uchiyama S., Vandenbol M., Vannier F., Vassaretli A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Welzenegger T.,
 RA Winters P., Wiput A., Yamamoto H., Yamane K., Yatsunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256 (1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSA SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D50543; BAA09107.1; -
 CC EMBL; Z59108; CAB12651.1; -

DR HSSP; P13569; 1MBD.
 DR Subtilisin; Bg11850; Yf1C.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transport.
 DR InterPro; IPR01140; ABCtransport.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR ProDom; PD00006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR Hypothetical protein; ATP-binding; Transport; Transmembrane;
 KM Complete proteome.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 172 192 POTENTIAL.
 FT TRANSMEM 288 308 POTENTIAL.
 FT TRANSMEM 510 530 POTENTIAL.
 FT NP_BIND 399 406 ATP (POTENTIAL).
 SQ SEQUENCE 604 AA; 67396 MW; 1998E469A2E6133 CRC64;

 Alignment Scores:
 Pred. No.: 1,12e-38 Length: 604
 Score: 674.50 Matches: 190
 Percent Similarity: 49.52% Conservative: 120
 Best Local Similarity: 30.35% Mismatches: 257
 Query Match: 15.86% Indels: 59
 DB: 1 Gaps: 16

 AF133659 (1-2345) x YF1C_BACSV (1-604)
 QY 310 CTCACACAGACCCAAAGAGGTTAAAGATGTGATCTCGAATAATCATAAAGCA 369
 DB 14 LeuProileAspLysLysGluGluGlyAlaLysLysArgAlaLysAlaLysAspThrLysGly 33
 QY 370 ATGCTTTCTTATGTGTGCC-----AAAGACAGCCAGATCTACAGACTGAGTTGCC 423
 DB 34 ThrLeuArgArgLysIleTrpSerTyrLeuAlaGluArgLysGlyLeuLeuIleLeuValMet 53
 QY 424 ATTGCGGCG-----GGATTTTGGGGGTGCGCAAGGCCATGAATATT 465
 DB 54 LeuMetValIleLysSerAlaIlePheGlyLeuLeuGly----- 66
 QY 466 GTGTTGCCCTTCATGTTAAATATGCTGACAGCCCTC-----AACGAGATGCG 516
 DB 67 -----ProPheValIleGlyLysAlaIleAspHisPheIleValGlyLysThrValSer 84
 QY 517 GGAACATGCTGAACCTGATGATGACCAATTAACATTGACCAATGCAACAGCAAGT 576
 DB 85 GlyLeuIle-----ProValLeuLeuLeuLeuAlaIle 96
 QY 577 CTGATTTGGCTATGTGTATCAGAGCTGAGTGTCTTTTAAAGAACTTGAAGTGA 636
 DB 97 TyrIleIleGlnSerLeuSer-----LeuTrpPhe-----GlnAsnTyr 109
 QY 637 GTATTGGCAAGTAGCCCAAGATTCAATCCGAGAAATAGCCAAATGCTTTCTTCAT 696
 DB 110 TrpMetIleThrIleSerGlnGlyThrValPheArgMetArgSerGlnLeuPheThrHis 129
 QY 697 CTTCACAACTGGATCTGGGTTTTCACCTGACAGACAGACAGCGAGCTTATCTAAGCT 756
 DB 130 LeuHisGluLeuProIleProPhePheAspLysGlnArgHisGlyGluLeuMetSerArg 149
 QY 757 ATTGACAGAGCAAAAGGGGTATCAGTTTGTCTGAGGCTTGGTATTTAATCTTCTT 816
 DB 150 ValThrAsnAspIleGluAsnValSerSerThrLeuAsnThrSerValIleGlnIleLeu 169
 QY 817 CCCATCATGTTTGAAGTAGTCTTGCAGTGTGTTTGTATTAATAATGCGGTGCCAG 876
 DB 170 SerSerValIleThrPheValGlyThrIleAlaValMetLeuTyr-----Met 185
 QY 877 TTTCGTTTGGTAACCTTGGAACTTGGTACATACAGCAATTCAAGTTCAGTGCAGTCA 936
 DB 186 SerProLeuLeuThrLeuIleThrLeu---ThrIleIleProValMetAlaIleSerLeu 204

QY 937 CGGTGAGAACTAGA-----TTTAGAATAGAAATGAACAAAGCAGATAAAT 981
 DB 205 LysTrpIleThrAsnArgThrGlyLysLeuPheLys---GluGlnGlnLysAsnLeuGly 223
 QY 982 GATCAGAGTAATGCTGCTATPAGACTCACTGCTGAATTAATGAACCTGGAAGTATTTAAT 1041
 DB 224 AspLeuasnGlyTyrIleGluGlnSerValSerGlyAlaLysValIleLysAlaTyrSer 243
 QY 1042 AATGAAGATATGAAGCAGAGATATGATGATTTTGAAGAGTATGAGACTGCTTCA 1101
 DB 244 ArgGluLys-----GlnIleThrAlaGluPheLeuGlu-----LysAsnAlaIle 258
 QY 1102 TTGAAAGTACCTGCTACTGCTGATGCTGAACCTTTGGT-----CAAAAGT 1146
 DB 259 LeuLysThrSerGlyPheThrAlaGlnThrIleSerGlyPheIleProLysValMetAsn 278
 QY 1147 GCTATTTCAGTGTGCTTAAACAGTAAATGCTGCTCCGACGTCAGGAAATTTGGCA 1206
 DB 279 SerLeuAsnLeuSerPheThrMetIleAlaIleGlyGlyLeuPheAlaLeuLys 298
 QY 1207 GGTACCTTACTGTGTGAGATCTAATGTAATGTAATGAGCTGCTTTTCAGCTTCAATTA 1266
 DB 299 GlyTrpIleSerIleGlySerIleValValPheAlaGluTyrSerArgGlnPheThrArg 318
 QY 1267 CCCCTGAACCTTCTGGGAACTGTATATAGAGAGACTAGACACAGCACTCATGATATGAC 1326
 DB 319 ProLeuAsnAspLeuAlaAsnGlnPheAsnThrMetLeuSerAlaIleAlaGlyAlaGlu 338
 QY 1337 ACCTGTTTACTCTATCTCAAGGTAGACACCCAAATTAAGCAAAAGTATGCTTCCC 1386
 DB 339 ArgValPheAspValIleAspGluLysGluGlnArgGlnAspGlyLysAsnAlaValHis 358
 QY 1387 CTTCAGATCACACACAGACAGCTACCGTGCTTTGATATGTCATTTGGAATACATT 1446
 DB 359 GlnProIle-----GlnThrGlySerIleGluPheArgAspValSerPheGlyTyrAsp 376
 QY 1447 GAGGCGCAAGAGTCTGTAGTGAATATCCTTGAAGTCCCTGGAGGAAAGAAAGTGGCC 1506
 DB 377 LysGlyGlnGlnThrLeuLysHisIleLeuGlnPheThrValProAlaGlnSerIleAla 396
 QY 1507 ATTTAGAGGATGAGTGGTCAAGGAAAGCAAAATAGTACGCTATTATTTGCTTTAT 1566
 DB 397 PheValGlyProThrGlnAlaGlyLysThrValThrAsnLeuLeuAlaArgPheTyr 416
 QY 1567 GAGCTCAAAAGGTAGATTTATCTTGCTGCTCAAAATATACAAATGACGCTGGA 1626
 DB 417 GluProAsnAspGlyLysIleLeuIleAspGlyThrAspIleLysThrLeuThrArgAla 436
 QY 1627 AGCTTGGAGGAGGAGTGGAGTGTACTCAGAGATGCTGCTTCATATATATTAATT 1686
 DB 437 SerLeuArgLysAsnMetGlyPheValLeuGlnHisPheLeuPheGlnGlyThrIle 456
 QY 1687 TATTACAACTTTATATAGAAATCATGCTTCACTGAGAGTGTATGACGTGGCA 1746
 DB 457 ArgGluAsnIleArgTyrGlyArgIleuAspAlaSerAspGlnGluValAlaAla 476
 QY 1747 AATTAGCTGACCTTCAATGATGCAATTTCTTGATGCCATGATATGACACCCAACTA 1806
 DB 477 LysThrAlaAsnAlaHisSerPheIleGluArgLeuProLysGlyTyrAspThrValLeu 496
 QY 1807 GGGGACAGGACTCAAGCTTTTCAGAGAGAGAAAGCAAAAGATGACAAATTCAGAGACC 1866
 DB 497 ThrGlnAsnGlySerGlyIleSerGlnGlyGlnLysGlnLeuIleSerIleAlaArgAla 516
 QY 1867 ATTTGAAGAGACCCCGCATCATCTATGATGAAGCTTCAATGCTTGAATTCGATT 1926
 DB 517 ValLeuAlaAspProValLeuLeuIleLeuAspGluAlaThrSerAsnIleAspThrVal 536
 QY 1927 ACTGAAGAGATATTTCTTGGTGCATGAGAGATGTGGTCAACACAGAACTTATTTTC 1986
 DB 537 ThrGluValAsnIleGlnGluAlaLeuAlaArgLeuMetGluGlyArgThrSerValIle 556

QY 1987 ATTGCACAGATGTTGACAGTGTGATGACAGATGAATCATTTGTTGATCAGAGGT 2046
DB 557 ILeAlahIsaNgLeuAaNTThrILeGlnArgAlaAspInILeValLeuLysAsnGly 576
QY 2047 AAGGTACCGCAACGTGTACCCACCATGTTGCTGCTTAACCTCATAGATCTATTCA 2106
DB 577 GluMetILeGluLysGlySerHisAspGluLeuLeu---ArgInLysGlyPheTyrSer 595
QY 2107 GAATGTGCATACACAG 2124
DB 596 AspleuTyrGluSerGln 601
RESULT 8
HLV2_ECOLI
ID HLV2_ECOLI STANDARD; PRT; 707 AA.
AC P10089;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion ATP-binding protein, chromosomal.
GN HLYB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J96 / Serotype O4;
RX MEDLINE=85234404; Pubmed=3891743;
RA Felmei T., Pellett S., Welch R.A.;
RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin";
RL J. Bacteriol. 163:94-105(1985).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M10133; AAA23976.1; -.
DR PIR; B24433; LEECB.
DR MEROPS; C39; UNK; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001140; ABC_transporter_TM.
DR InterPro; IPR005074; peptidase_C39.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF03412; peptidase_C39; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Hemolysis; Transport; ATP-binding; Transmembrane.
KT TRANSMEM 158 179 POTENTIAL.
FT TRANSMEM 187 204 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 293 311 POTENTIAL.
FT TRANSMEM 360 377 POTENTIAL.
FT TRANSMEM 387 403 POTENTIAL.
FT TRANSMEM 407 425 POTENTIAL.
FT TRANSMEM 425 509 POTENTIAL.
FT NE_BIND 502 509 ATP (BY SIMILARITY).
SQ SEQUENCE 707 AA; 79463 MW; 21009CB45E59437E CRC64;

Alignment Scores:
Pred. No.: 1,34e-38 Length: 707
Score: 673.50 Matches: 176
Percent Similarity: 51.69% Conservative: 115

Best Local Similarity: 31.26% Mismatches: 239
Query Match: 15.83% Indels: 33
DB: 1 Gaps: 11
AF133659 (1-2345) x HLY2_ECOLI (1-707)
QY 463 ATTGTGGTTCCTTCATGTTTAATATGCTGTGACAGCCTCAACAGATGTGGCA--- 519
DB 169 LeuILeThrProLeuPhePheGlnValValMetAspLysValLeuValHisArgGlyPhe 188
QY 520 AACATGCTGAACCTGATGATGACCAACAAATPACGTTCGAACGCAACAGCAGTTCTG 579
DB 189 SerThrLeuAsn-----ValILeThrValAlaLeuSerValVal 201
QY 580 ATTGCGTATGCTGATCAAGAGCTGAGCTGCTTTTAAACGAATTCGAATGACGTA 639
DB 202 ValValPheGluLeuLeuLeuSerGly-----LeuArgThrTyrIle 215
QY 640 TTTCGAAGTACCCAGAAATTCATCCGAAGATAGCCAAAATGCTTTTCATCTT 699
DB 216 PheAlaHisSerThrSerArgIleAspValGluLeuGlyValAlaLysLeuPheArgHisLeu 235
QY 700 CACAACCTGATGTGGCTTTTACCTGACGACGACAGCAGGAGCTTATTAAGCTATT 759
DB 236 LeuAlaLeuProILeSerThrPheGluSerArgArgValGlyAspThrValAlaArgVal 255
QY 760 GACAGAGAAACAAGGGTATCAGTTTGTCTGAGGCTTTGGTATTTAATCTTCCC 819
DB 256 ArgGluLeuAspGlnILeArgAsnPheLeuThrGlyGlnAlaLeuThrSerValLeuAsp 275
QY 820 ATCATGTTGAAGTATGCTGTGTGACGTGCTTTGTATTAACAATCCGTCAGTTT 879
DB 276 LeuLeuPheSerPheILePhePheAlaValMetTyrTyr-----SerProLysLeu 293
QY 880 GCTTGTGTAACCTTGGAACCTGTGTATACATACAGCAAGCTTCAAGTGTGACGACA--- 936
DB 294 ThrLeuValILeLeuPheSerLeuProCysTyrAlaAlaThrSerValPheILeSerPro 313
QY 937 -----CGGTGAGAACTAGATTAGAAATGAATGAACAAAGCAGATATGACAGT 990
DB 314 ILeLeuArgArgArgLeuAspAspLysPheSerArgAsn---AlaAspAsnGln----- 330
QY 991 AATGCTGATAGACTCACTGCTGAATTGAACCTGGAATGATTATTAATAAGAAAGA 1050
DB 331 SerPheLeuValGluSerValThrAlaIleAsnThrILeLysAlaMetAlaValSerPro 350
QY 1051 TATGAAGCACAGAGATATGATGATTTTGAAGACGTATGACGCTTCATTGAAGAAAGT 1110
DB 351 GluMetThrAsnILeThrAspLysGlnLeuAlaGlyTyrValAlaIaGlyPheLysVal 370
QY 1111 ACCTGACTGTGCTATGCGTGAACCTTGTGCAAGGCTATTTTCAGTGTGCGTTTAACA 1170
DB 371 Thr-----ValLeuAlaThrILeGlyGlnGlnGlyILeGlnLeuILeGlnLysThr 387
QY 1171 GCTATATAGTG------CTCGCAGCTCAGGAATGTGGCAGTACCTTACTGTT 1221
DB 388 ValMetILeILeAsnLeuThrLeuGlyAlaHisLeuValILeSerGlyLysLeuSerIle 407
QY 1222 GGAAGATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1281
DB 408 GlyGlnLeuILeLysPheAsnMetLeuAlaGlyGlnILeValAlaIaProValILeArgLeu 427
QY 1282 GGAAGTATATATGAGAGCTAGCAAGCAAGCATCATGATATGAAACCTGTATTACTCTA 1341
DB 428 AlaGlnILeThrPheILeAspPheGlnGlnValGlyILeSerValThrArgLeuGlyAspVal 447
QY 1342 CTCAGATGACACCCAAATTAAGCAAGATGATGATGATGATGATGATGATGATGATG 1401
DB 448 LeuAsnSerProThrLysSerTyrHisGlyLysLeuAlaLeuPro-----Glu 463
QY 1402 CAGACAGCTACCGTGGCTTTGATATGTCATTTGAATAC---ATTAGGCGCAGAAA 1458
DB 464 ILeAsnGlyAspILeThrPheArgAsnILeArgPheArgTyrLysProAspSerProVal 483

Oy	1459	GTCCTTACTGGAATAATCCCTTTGAAGTCGCCCGCAGGAAAGAAAAGTGCCATTGTACAGAGT	1518
Db	484	IleLeuAspAsnIleAsnLeuSerIleYelgInglYglValalIlelYlIleValIglYarg	503
Oy	1519	AGTGGGTGAGGAAAAGACACAATGAGGCTATTAATTTGCTTTATGAGGCTCAAGAAG	1578
Db	504	SerGlySerGlyLysSerThrLeuThrLeuIleGlNArgHethrYlleProGluNsn	523
Oy	1579	GGTAGCATTTATCTTGCTGGTGCMAAATATACAAAGATGTAGGCTGGAAAGCCTTGGAGG	1638
Db	524	GlyAlnValIleuIleAspGlyHIsAspLeuAlaIleuAlaAspProAsnTrpLeuArgarg	543
Oy	1639	GCAGGGGAGGGTACCTCCAGGATGCTGTGCTCTCCATCAATATCTTTATTAACAACCTC	1698
Db	544	GlnValIglYvalIleuGlnAspAsnValIleuLeuAsnArgserIleIleAspSnlIe	563
Oy	1699	TTATATGGAACATCAGTGCTTCACCTGAGAAAGTGATGACAGTGCCAAMATTAGCTGGA	1758
Db	564	SerLeuAlaAsnProGlyMetSerValIglulYValIlleTyZalaIlaIylsLeuAlagly	583
Oy	1759	CTTCATATGCAATTTCTTCGAATGCCACATGATATGACACCACAATAGGAGGACGAGA	1818
Db	584	AlahIsaSpPheIleSerGluLeuArgIglulYTyrAsnThrIleValIglYlunGlngly	603
Oy	1819	CTCAAGCTTTTAGGAGGAGAAACAAAGAAGTGACATTCGCAAGGACCTTTTGAAGAC	1878
Db	604	AlaGlyLeuSerGlyGlyArgIunGlnArgIleAlaIleAlaArgAlaIleuValIasnSn	623
Oy	1879	CCCCCAGTCACTACTATATGAGAAGACTACTTCATCGTTAGATTGCATTACTGAGAGACT	1938
Db	624	ProlYslIleuIlePheAspGluAlaThrSerAlaIeuAspPyrgIusergIuhIsvaI	643
Oy	1939	ATTCTTGTCGCATGAGAGATGTGTGTAACAACAGAACCTTATTTTCAATTCACACAGA	1998
Db	644	IleMetArGaAsnMetHIsLysIleCySLysGlyArgThrValIleIleIleAlahIsarg	663
Oy	1999	TGTCAACAGAGGTGTGATGCAGATGAAATCATTTGTCTGGATCAGGGTACAGGTACACAA	2058
Db	664	LeuSerThrValIylsAsnAlaAspArgIleIleValaMecIulubSelyslYsrIleValGlu	683
Oy	2059	CGTGTACCCACCAACGATGGTTGCTGCTTAACCTTCATAGTATGTATTCAGAAATGTGGCAT	2118
Db	684	GlnglyLshIsLysGluLeuLeuSerGluProGluSerLeuYrseryrLeuYrcIn	703
Oy	2119	ACACAGAGC 2127	
Db	704	LeuGlnSer 706	
RESULT 9			
HLXB_ECOLI			
ID	_HLXB_ECOLI	STANDARD;	PRT; 707 AA.
AC	P08716;		
DT	01-JAN-1988 (Rel. 06, Created)		
DT	01-JAN-1988 (Rel. 06, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hemolysin secretion ATP-binding protein, plasmid.		
GN	HLXB.		
OS	Escherichia coli.		
OC	Plasmid Inc12 pHLV152.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia.		
OX	NCBI_TaxId=562;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Hess J., Wells W., Vogel M., Goebel W.;		
RT	"Nucleotide sequence of a plasmid-encoded hemolysin determinant and		
RT	its comparison with a corresponding chromosomal hemolysin sequence."		
RL	FEMS Microbiol. Lett. 34:1-11(1986).		
RJ	[2]		
RP	TOPOLOGY.		
KX	MEDLINE=92204133; PubMed=1552901;		
RA	Genlschev I., Goebel W.;		

```

RT      "Topological and functional studies on HlyB of Escherichia coli.",
RM      Mol. Gen. Genet. 232:40-48(1992).
RN      [3]
RP      TOPOLOGY.
RX      MEDLINE=91132653; PubMed=1994034;
RA      Wang R.C., Serror S.J., Blight M., Pratt J.M., Broome-Smith J.K.,
RA      Holland J.B.
RT      "Analysis of the membrane organization of an Escherichia coli protein
RT      translocator, HlyB, a member of a large family of prokaryote and
RT      eukaryote surface transport proteins."
RL      J. Mol. Biol. 217:441-454(1991).
CC      1. - FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
CC      1. - SUBUNIT: HOMODIMER.
CC      1. - SUBCELLULAR LOCATION: Integral membrane protein.
CC      1. - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILYX.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collabor
CC      between the Swiss Institute of Bioinformatics and the EMBL outstatc
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no
CC      modified and this statement is not removed. Usage by and for commec
CC      entities requires a license agreement (See http://www.isb-sib.ch/anno
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL, M14107, AAA98234.1, -.
DR      PIR, S10057, S10057.
DR      MEROPS, C39, UNW, -.
DR      InterPro: IPR003593; AAA ATPase.
DR      InterPro: IPR003439; ABC transporter.
DR      InterPro: IPR001140; ABCtransportTM.
DR      InterPro: IPR005074; peptidase_C39.
DR      Pfam, PF00005; ABC_tran, 1.
DR      Pfam, PF00664; ABC_membrane, 1.
DR      Pfam, PF03412; Peptidase_C39, 1.
DR      Prodom: PD000006; ABC_transporter, 1.
DR      SMART: SM00382; AAA, 1.
DR      PROSITE, PS00211; ABC_TRANSPORTER, 1.
DR      HemoIySis; transporter; ATP-binding; Transmembrane; Plasmid.
KM      HemoIySis; transporter; ATP-binding; Transmembrane; Plasmid.
FT      DOMAIN 1 37
FT      TRANSMEM 38 51
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 52 57
FT      TRANSMEM 58 78
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 79 157
FT      TRANSMEM 158 179
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 180 190
FT      TRANSMEM 191 211
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 212 268
FT      TRANSMEM 269 289
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 290 294
FT      TRANSMEM 295 315
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 316 387
FT      TRANSMEM 388 408
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 409 411
FT      TRANSMEM 412 432
FT      PERIPLASMIC (PROBABLE).
FT      DOMAIN 433 707
FT      TRANSMEM 433 707
FT      PERIPLASMIC (PROBABLE).
FT      NP BIND 502 509
FT      ATP (BY SIMILARITY).
SQ      SEQUENCE 707 AA; 79672 MW; 412A3EB64A3CFBA CRC64;

Alignment Scores:
Pred. No.: 1,83e-38 Length: 707
Score: 671.50 Matches: 175
Percent Similarity: 51.69% Conservative: 116
Best Local Similarity: 31.08% Mismatches: 239
Query Match: 15.79% Indels: 33
DB: 1 Gaps: 11

AF133659 (1-2345) x HLYB_ECOLI (1-707)
QY 463 ATTTGGATCCCTTCATGTTAATATGCTGTAGACAGCCTCAACGATGTCGGG--- 519
DB 169 LeuIleIleMrProLeuPhePheGlnValValMetAspIysrValIleuValIhIeAggIyPhe 188
QY 520 AACATGCTGGAACCTGAGTGATGACCAAAATATACAGTTGCCAACATGCGACACAGAGTTCTG 579

```

```

Db 189 SerThrLeuAen-----ValIleThrValAlaLeuSerVal 201
QY 580 ATTGGCTATGGTATCATCAAGAGCTGGACCTGTTTAAAGCAATGCAAGTGA 639
Db 202 ValValPheGluIleIleLeuSerGly-----LeuArgThrTyrIle 215
QY 640 TTGGCAAGGATGACCCAGAAATTCAGTCCGAAATAGCCAAAATGCTTTCTCATCTT 699
Db 216 PheAlaHisSerThrSerArgIleAspValGluLeuGlyAlaIleValSerPheArgHisLeu 235
QY 700 CACAACTCGATCTGGGTTTACCTGAGCAGACAGGAGCTTATCTTAAGCTATT 759
Db 236 LeuAlaLeuProIleSerTyrPheGluSerArgValGlyAspThrValAlaArgVal 255
QY 760 GACAGAGAAACAAGGGATATGCTTGTCTGAGTGCTTGATTTAATCTTCCCTCC 819
Db 256 ArgGluLeuAspGlnIleArgAsnPheLeuThrGlyAlaIleLeuThrSerValLeuAsp 275
QY 820 ATCATGTTTGAAGTATGATGCTTGTGAGTGCTTTGATTAACAAATGCGGTGCCAGTTT 879
Db 276 LeuLeuPheSerLeuIlePhePheAlaValMetTrpTyr-----SerProGlyLeu 293
QY 880 GCTTGGTAACTTGGCAACCTTGGTATCATACAGCAGATTACAGTTGACAGTACA--- 936
Db 294 ThrLeuValIleLeuPheSerLeuProCysTyrAlaIleThrSerValPheIleSerPro 313
QY 937 -----CGGTGAGACACTAGATTGAATGAATGAACAAGCAAGTAATGATGCAAGT 990
Db 314 IleLeuArgThrArgLeuAspAspLeuPheSerArgAsn--AlaAspAsnGln----- 330
QY 991 AATCTGCTATAGACTGACTGCTGAATTAAGCACTGAGATTTTAAATATGAAGA 1050
Db 331 SerPheLeuValGluSerValThrAlaIleAsnThrIleGlyAlaMetAlaValSerPro 350
QY 1051 TATGAGACACAGAGATATGATGATTTTGAAGACGATGAGACTGCTTCAATGAAGA 1110
Db 351 GlnMetThrAsnIleTrpAspLeuGlnLeuAlaGlyTyrValAlaIleAlaGlyPheGlyVal 370
QY 1111 ACCCTGACTGCTGCTATGCTGAACCTTGGTCAAGTCAATTTTCAGTTCGGTTTACA 1170
Db 371 Thr-----ValLeuAlaThrIleGlyGlnGlnGlyIleGlnLeuIleGlnLeuYthr 387
QY 1171 GCTATAATGCTG-----CTGCCAGTCAAGGAAATTTGCGAGGTACCTTACTGTT 1221
Db 388 ValMetIleIleAsnLeuTrpLeuGlyAlaHisLeuValIleSerGlyAspLeuSerIle 407
QY 1222 GGAGATCTTATGATGATGATGAGCTGCTCTTTTACGCTTCATTATCCCTGAACTTTCTG 1281
Db 408 GlyGlnLeuIleAlaPheAsnMetLeuAlaGlyGlnIleValAlaProValIleArgLeu 427
QY 1282 GGAAGCTGATATGAGAGACTAGACAGACACTCATATATGAAACACTTGTACTCTA 1341
Db 428 AlaglnIleTrpGlnAspPheGlnGlnValGlyIleSerValThrArgLeuGlyAspVal 447
QY 1342 CTCAGGTAGACACCCAAATTAAGCAAAAGTGAATGATGCCATCCCTTCAGATCACCA 1401
Db 448 LeuAsnSerProThrGlu-----SerTyrHisGlyGlySerLeuThrLeuProGlu 463
QY 1402 CAGACAGCTACCGTGGCTTGTATATGTCATTTGAATAC--ATTGAGGGCCAGAAA 1458
Db 464 IleAsnGlyAspIleThrPheArgAsnIleArgPheArgTyrAspProAspSerProVal 483
QY 1459 GTCCTTATGGAATATCTTGAAGTCCCTGCGAGAAAGAAAGGCGCATTTAGAGAGT 1518
Db 484 IleLeuAspAsnIleAsnLeuSerIleGlyGlnGlnValIleGlyIleValGlyArg 503
QY 1519 AGTGGCTCAGGAGAAACACAATATGATGAGGCTATTATTTGCTTCTTATGAGCCCAAG 1578
Db 504 SerGlySerGlyGlySerThrLeuThrValLeuIleGlnArgPheTyrIleProGlnAsn 523
QY 1579 GGTAGCATTTATCTTCTGCTGCTCAAAATATATCAAGATGTGAGCTTGAAGACCTTGGAGG 1638

```

```

Db 524 GlyGlnValLeuIleAspGlyHisAspLeuAlaLeuAlaAspProAsnTrpLeuArgArg 543
QY 1639 GCAGTGGAGATGGTGAACCTCAGATGCTGCTCTTCATATATCTATTATTACAACTC 1698
Db 544 GlnValGlyValValLeuGlnAspAsnValLeuLeuAsnArgSerIleIleAspAsnIle 563
QY 1699 TTATATGAAACATCACTAGCTGCTTCACTGAGGAAGTGTATGCACTGGCAAAATTAAGTGA 1758
Db 564 SerLeuAlaAsnProGlyMetSerValGluLeuValIleTyrAlaIleValSerLeuAlaGly 583
QY 1759 CTTTATGATGATCAATTTCTTGAATGCCATGATATGACACCCCAAGTAGGGGAACGAGA 1818
Db 584 AlaHisAspPheIleSerGluLeuArgGluGlyTyrAsnThrIleValGlyGluGlnGly 603
QY 1819 CTCAGGCTTCAAGAGAGAAAGCAAAAGAGTGAATTCAGAGCCATTTTGAAGAGAC 1878
Db 604 AlaGlyLeuSerGlyGlyGlnArgGlnArgIleAlaIleAlaArgAlaLeuValAsnAsn 623
QY 1879 CCCCAGTCACTACTTATGATGAGACTTACTTCACTGATGATGATGATGATGATGATGAT 1938
Db 624 ProIleIleLeuIlePheAspGlnAlaThrSerAlaLeuAspTyrGluSerGluHisVal 643
QY 1939 ATTCTTGCTGCCATGAAGAGTGTGCTCAACACAGAACTTCTATTTCATTCAGACACAGA 1998
Db 644 IleMetArgAsnMetHisIleCysGlyArgTyrValIleIleIleAlaHisArg 663
QY 1999 TTGTCAACAGTGGTTGATGAGATGAAATCATGCTTGGATGAGGTAGGTAGGCCGA 2058
Db 664 LeuSerThrValIleAsnAlaAspArgIleIleValMetCysGlyGlyValGlu 683
QY 2059 CGTGTATACCACCATGCTGTTGCTTGGTAACTTCATATGATCTTATTCAGAAATGTGAGAT 2118
Db 684 GlnGlyLysHisIleGlyGluLeuLeuSerGluProGluSerLeuTyrSerTyrLeuTyrGln 703
QY 2119 ACAAGAGC 2127
Db 704 LeuGlnSer 706

RESULT 10
AB10_MOUSE STANFORD; PRT; 715 AA.
ID AB10_MOUSE Q9J139; Q9J0C7;
AC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP-binding cassette, sub-family B, member 10, mitochondrial precursor
DE (ATP-binding cassette transpporter 10) (ABC transpporter 10 protein)
DE (ABC-mitochondrial erythroid protein) (ABC-me protein).
GN ABCB10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Erythrocyte;
RX PubMed=10835348;
RA Shirai O.S., Gregory T., Yu C., Ozkin S.H., Weiss M.J.;
RT "ABC-me: a novel mitochondrial transporter induced by GATA-1 during
RT erythroid differentiation.";
RL EMBO J. 19:2492-2502(2000).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saio T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saio R.,
RA Kadota K., Matsuda H.A., Ashburner M., Betalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikiyama T., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

```



```

Db 304 -----leuserlysalatThrGlnAspSerLeuAlaGluAlaThrGlnLeuAlaG1 320
Qy 1002 AGACTCACTGCTGAATTAATGAAACCTGCAAGTATTTAAATGAAGAAATGAAGACA 1061
    ::::: ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 320 uGUAArgIleGlyAsnIleArgThrIleArgAlaPheGlyLysGlnMetThrGlnValG1 340
Qy 1062 GAGATATGATGGATTTTGAAGACGATGAGACGCTTCATTGAAAGTAAAGTCTTACTCT 1121
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 340 uUyStyTrhGly-----ArgValAspGlnLeuGlnLeuAlaGlnLysGlnValaLe 358
Qy 1122 GGCATATGCTGAACCTTGGTCAAGTCATTTTCAGTGTGGCTTAAACGCTAATATAGCT 1181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 uAlaArgAlaGlyPhePheGlyAlaAlaGlyLeuSerGlyAsnLeuIleValLeuSerVa 378
Qy 1182 GCTGCCAGTCAGGGAATTTGCGAGGTACC-----CTTACTGTTGAGATCTAGTAAT 1235
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 378 lLeuTrpLysGlyGlyLeuLeuMetGlySerAlaHisMetThrValGlyGlnLeuSerSe 398
Qy 1236 GGTGAATGACTGCTTTTTCAGCTTTCATTACCCCTGAACCTTTGCGAATGTATATAG 1295
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 398 rPheLeuMetTrpAlaPheTrpValGlyLeuSerIleGlyLysSerSerPheTrpYse 418
Qy 1296 AGAGACTAGACACACTCATAGATATGAAACCTTGTTCATCTACTCAAGTAGACAC 1355
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 418 rGlnLeuMetLysGlyLeuGlyAlaGlyArgLeuTrpGlnLeuGlnArgGlnPr 438
Qy 1356 CCAA-----ATTAAGACAAATGATGGATGCCCTCCCTCA 1391
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 438 oArgLeuProPheAsnGlnGlyMetValLeuAspGlnLysThrPheGlnGlyAlaLeuG1 458
Qy 1392 GATCACACACAGACACTACCGTGGCTTTGATATATGCAATTTTGAATATCATTTGAGG 1451
    ::||| ||||| ||||| ||||| ||||| ||||| |||||
Db 458 u-----PheArgAsnValHisPheThrTrpProAlaAr 469
Qy 1452 CCAGAA-----CTCCTTAGTGAATATCTTTGAAGTCCCTGCAGAAAGAAAGTGGC 1505
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 469 gProGlnValSerValPheGlnAspPheSerSerIleProSerGlySerValThrAl 489
Qy 1506 CATTTGAGAGTGGTGGTCAAGGAAAGACAAATAGTGGCTTATATTTGCTCTTA 1565
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 aLeuValGlyProSerGlySerGlyLysSerThrValValSerLeuLeuAspGlnLeuTr 509
Qy 1566 TGAAGCTCAAAAGGTAGCATTTATCTGTGTCGTAACAAATATACAGATGTAGCCTGGA 1625
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 509 rAspProAsnSerGlyThrValSerLeuAspGlyHisAspIleArgGlnLeuAsnProVa 529
Qy 1626 AAGCTTCGAGAGGCGAGTGGAGTGTACTCAGAGATGCTGCTCTTCATTAATCTAT 1685
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 529 lTrpLeuArgSerLysIleGlyThrValSerGlnGluProValLeuPheSerCysSerVa 549
Qy 1686 TTATTATCACTCTTATATGGA-----AACATCAGTGC---TCACCTAGAGAAAGTGA 1736
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 549 lAlaGlnAsnIleAlaTrpGlyAlaAspAsnLeuSerSerValThrAlaGlnValG1 569
Qy 1737 TGCAGTGGCAAAATTAAGTCTGACTTCATGATGCAATCTTCGATGACATGCGATATGA 1796
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 569 uArgAlaAlaGlnValAlaAsnAlaAlaGlnPheIleArgSerPheProGlnGlyPheAs 589
Qy 1797 CACCAAGTGGGGAACAGAGACTCAAGCTTTTCAGAGAGAGAAAGCAAGAGTAGCAAT 1856
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 589 pThrValValGlyGlnLysGlyIleLeuLeuSerGlyGlyGlnLysGlnArgIleAlaI1 609
Qy 1857 TGAAGAGCACTTTTGAAGAGCCCAAGTCATCTATAGTAAAGTCAACTTCATCGTT 1916
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 609 eAlaArgAlaLeuLeuLysAsnProLysIleLeuLeuLeuAspIleAlaThrSerAlaLe 629
Qy 1917 AGATTGATTACTGAAGAGACTATTTCTGGTGCATGAGAGATGCTCAACACAGAAC 1976
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 629 uAspAlaGlnLysGlnIleuValGlnGlnAlaLeuAspArgLeuMetGlnGlyArgTh 649
Qy 1977 TTCTATTTCATTGACACAGATTGTCACAGTGGTTGATGACAGATGAAATCATTTGCTT 2036
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||

```

```

Db 649 rValLeuIleIleAlaHisArgLeuSerThrIleLysAsnAlaAsnPheValAlaValLe 669
Qy 2037 GATACAGGTAAGTAAAGCCGAAACGTGTACCCACAGTGTCTTCTGCTCACTCATAG 2096
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 669 uAspHisGlyLysIleLeuGlnLysIleuThrHisGlnGlnLeuLeuLysProAsnG1 689
Qy 2097 TATCTATTCAGAAATGTGGCATACACAGAC 2127
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db 689 yLeuTrpArgLysLeuMetAsnLysGlnSer 699
RESULT 11
HLVB_PASHA STANDARD; PRT; 708 AA.
ID HLVB_PASHA
AC P16532;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukotoxin secretion ATP-binding protein.
GN LKTB.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype A1;
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella
    haemolytica A1.";
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype A1;
RX MEDLINE=89123172; PubMed=2914876;
RA Strathdee C.A., Lo R.Y.C.;
RT "Cloning, nucleotide sequence, and characterization of genes encoding
    the secretion function of the Pasteurella haemolytica leukotoxin
    determinant.";
RL J. Bacteriol. 171:916-928(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype A1 / PH101;
RX MEDLINE=89210283; PubMed=2707120;
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
    cluster.";
RL DNA 8:15-28(1989).
RN [4]
RP SEQUENCE OF 1-48 FROM N.A.
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstock G.M.;
RT "Secretion and expression of the Pasteurella haemolytica leukotoxin.";
RL J. Bacteriol. 172:2343-2350(1990).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLVB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M20730; AAA25530.1; -.
DR EMBL; M24197; AAA25544.1; -.
DR PIR; A32051; A32051.
DR PIR; S29517; S29517.
DR MEROPS; C39.UNW; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.

```


D 534 LeuAlaLeuAlaLeuAaPProAenTTPLeuAArgGlnIleGlyValValLeuGlnAAspAsn 553
Q 1666 GTCCCTTCATTAATACATATTATTAACAACCTCTTAATAGGAACATCAGTGTCTCACT 1725
D 554 ValLeuLeuAenAArgSerIleAArgGlnAsnIleAlaLeuSerAAspProGlyMetProMet 573
Q 1726 GAGGAAGTGTATGCAAGTGGCAAAATTAGCTGGACTTCATGATGCAATTCTTTCGAATGCCA 1785
D 574 GluArgValIleTyrAlaIleAlaValLeuAlaGlyAlaIleAAspPheIleSerGlnLeuArg 593
Q 1786 CATGATATATGACACCCCAAGTGGGGAAGAGAGACTCAAGCTTTTACAGAGAAAGCA 1845
D 594 GluGlyTyrAsnThrIleValGlyGlnGlnIleValaGlyLeuSerGlyGlyGlnArgGln 613
Q 1846 AGATAGCAATATGCAAGAGCCATTTTGAAGACCCCAAGTCACTATCTATGATGAAGCT 1905
D 614 ArgIleAlaIleAlaArgAlaLeuValAAsnAsnProGlyIleLeuIlePheAspGlnAla 633
Q 1906 ACTTCATGCTTGAATTCGATTACTGAAAGACATATCTTGGTCCATGAAGATGTGTC 1965
D 634 ThrSerAlaLeuAAspTyrGlnSerGlnIleIleIleMetGlnAAsnMetGlnIleCys 653
Q 1966 AAACACGAGACTTCTATTTTCATTCGACACAGATTCACAGTGTGATGCAAGTGA 2025
D 654 GlnGlyTyrThrValIleLeuIleAlaIleAAspSerThrValIlyAsnAlaAAspArg 673
Q 2026 ATCATGCTTGTGATGAGGATGAGTACCCAGTGTGATCCCAAGTGTGCTTCTGCT 2085
D 674 IleIleValMetGlnGlySerGlyGlnIleValGlnGlnIlySerIleGlnLeuLeuGln 693
Q 2086 AACCTCATAGTATCTATTCAGAAATGTGGATACAG 2124
D 694 AsnSerAAsnGlyLeuTyrSerTyrLeuHISGlnLeuGln 706

RESULT 12

Y4GM_RHISN STANDARD; PRT; 586 AA.
ID Y4GM_RHISN
AC P55469;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ABC transporter ATP-binding protein Y4GM.
GN Y4GM.
OS Rhizobium sp. (strain NGR234) .
OC Plasmid sym PNGR234.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
ON NCBI_taxid=394;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
Berret X.,
RT Molecular basis of symbiosis between Rhizobium and legumes."
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY, MSBA SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AB000075; AAB91687.1; .
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; AAA_ATPase.
CC InterPro; IPR001140; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00664; ABC_membrane; 1.
CC ProDom; PD000006; ABC_transport; 1.
CC SMART; SM00382; AAA; 1.

DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; ATP-binding; Transport; Transmembrane; Plasmid.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT NP_BIND 379 386 ATP (POTENTIAL).
SQ SEQUENCE 586 AA; 64262 MW; 1095DFB82620637 CRC64;
Alignment Scores:
Pred. No.: 2,82e-37 Length: 586
Score: 654.00 Matches: 175
Percent Similarity: 51.59% Conservative: 100
Best Local Similarity: 32.83% Mismatches: 226
Query Match: 15.37% Indels: 32
DB: 1 Gaps: 13

AF133659 (1-2345) x Y4GM_RHISN (1-586)

Q 562 ATGCAACAGACAGTTCGATGCTGATGCTATCAAGACCTGAGCTGTTTTCAC 621
D 69 ValAlaValThrValAlaIleIlePheAlaValIle-----GlyLeuAlaThrTyr--- 85
Q 622 GAAGTTCGAATGACAGTATTGGCAAGTACCCAGATTCATCCGAGAATAGCCAA 681
D 86 ---ValGlnSerIlePheLeuSerIlyAlaGlyAAsnIleIleAlaIsthrGlnArg 104
Q 682 AATGCTTCCTTCATCTTCACACCTGATCTGGGTTTTCACCTGACACAGACGGA 741
D 105 ArgLeuPheGlnIlyIsthrValLeuArgGlnGlyLeuSerPheTyrIleTyrProSerSer 124
Q 742 GCTTTATCTAAGGCTATTGTGACAGAGAACAGGAGTATCATGTTTGCTCGAGTCTTG 801
D 125 GluLeuLeuValArgLeuThrAAsnAlaGlnAlaValArgSerValIleGlnLeuVal 144
Q 802 GATTTAATCTTCCATCATGTTTGAAGTATGATGCTGTGATGCTGTTTGTATTAC 861
D 145 ValThrSerPheIleAArgAAspLeuPheSerLeuMetGlyLeuLeuAlaValMetVal--- 163
Q 862 AATGCGGTCGCCAGTTGCTTGGTGAACCTT-----GGAACACTGTTTACATACACA 915
D 164 -----IleGlnIleProLeuLeuSerLeuValSerIleAlaValGlyProGlyAla 180
Q 916 GCATTCACAGTTCACAGTACACGCTGAGAACATGATTTAGATTAAGAACAAAGCA 975
D 181 IleLeuGlyValArgValLeuThrArgIysValArgIlyIleMetGlnLeuGlnIleAla 200
Q 976 GATTAATGACAGATATGCTGCTATAGACTCACTGCTGAATTATGAACT----- 1026
D 201 -----SerIleGlyGlnIleIleGlnSerValGlnIleThrSerThrGlyIleArg 217
Q 1027 ---GTGAAGATTTTAAATATGAAGAATATGACACAGATATGATGATTTTGAAG 1083
D 218 IleValIlyIsthrAlaLeuGlnAAspPheMetArgArgIsthrAspIlyTyrIleGly 237
Q 1084 ACGTATGAG-----ACGCTTCATTAAGAAAGTCTCTACTCTGCGGTATG 1128
D 238 AspValGlnArgArgAlaAsnSerIleAlaArgLeuGlnAlaIleAspSerProIleMet 257
Q 1129 CTGAACCTTGGTCAAGTGTATTTTCAGTGTG---GGTTTAACACTAATATGAGTCTC 1185
D 258 GluThrLeuSerGlyPheAlaIleAlaGlyValIleAlaLeuSerGlyValLeuValLeu 277
Q 1186 GCGAGTCAGGAATTTGTGGCAGGTACCTTACTGTTGGAGATCTAGTAATGTGAATGCA 1245
D 278 ---GlnGlnGlyAAsnThrProGlyGlnLeu-----MetSerPheIleThrAla 292
Q 1246 CTGCTTTTCAGCTTCATTAACCTGGAACCTTTCGGAACCTTAATATAGAGAGACTAGA 1305
D 293 LeuLeuLeuAlaTyrGln---ProAlaIlyArgLeuAlaArgMetCArgIleSerLeuGln 311
Q 1306 CAAGCACTATAGATATGAACACTTGTATTACTATCTCAAGGTAGACACCAAAATTAA 1365

Db 312 SerAlaLeuValGlyValArgMetMetYrGlnLeuAlaAspHisProIleGluLeuThr 331
 QY 1366 GACAAAGATGATGGATCCCTTCAGATCACACAGACAGATCCAGTGGCCCTTGAT 1425
 Db 332 GlnValAsnSerAlaIleProLeuProGluGlyPro-----GlyGlnIleArgPheLeu 349
 QY 1426 AATGTGATTTTGAATATGATTCAGGAGCCAGAAAGTCTTATGTAATTCCTTTGAGATC 1485
 Db 350 AspValAsnProSerTyrLysAsnGlyGlnArgLeuPheGlnAsnLeuAsnValThrPhe 369
 QY 1486 CTGCGAGAAAGAAAGTGGCCATTGTAGAGGTAGTGGTTCAGGAGAAAGCAATAGTGTG 1545
 Db 370 ProAlaGlyLysThrThrAlaLeuValGlyProSerGlyAlaGlyLysSerSerIleIle 389
 QY 1546 AGGCAATATTTCGCTTCATGAGCTCAAAAGGTAGATTTATCTTCGCTGCAAT 1605
 Db 390 AsnLeuIleMetArgLeuTyrAspProAspValGlySerValThrValAspGlyHisAsp 409
 QY 1606 ATACAGATGTGAGCCCTGAGAAAGCTTCGAGGAGGAGTGGAGTGGTCTCAGATGCT 1665
 Db 410 LeuLysAspValThrPheArgSerLeuThrGAspArgGlyIleGlyPheValGlyGlnAspThr 429
 QY 1666 GTCTCTTCATATATGATTTATTAACAACCTTTATATGAAACATAGTGCCTTCACT 1725
 Db 430 PheLeuPheSerGlyThrIleLysTyrAsnIleSerLeuGlyArgGluGlyAlaSerAsp 449
 QY 1726 GAGGAGAGTATGAGTGGCAAAATTAGTGCATCTTCATGATGCAATCTTCGATGCGCA 1785
 Db 450 GlnGluIleIleGluAlaAlaLysThrAlaAsnAlaHisAspPheIleMetLysMetPro 469
 QY 1786 CATGATATGACACCCCAATAGGAGGAGACGAGATCTCAAGCTTTGAGAGAGAAAGCAA 1845
 Db 470 HisGlyTyrAspThrGluValGlyGlnAsnGlyIleLysLeuSerGlyGlyGlnLysGln 489
 QY 1846 AGAGTAGAATGCAAGAGCCATTGTAAGAGACCCCAATCATCTATGATGATGCT 1905
 Db 490 ArgIleThrIleAlaArgAlaMetLeuArgAsnAlaGluIleLeuIlePheAspGluAla 509
 QY 1906 ACTTCATGTTAGTATGATTCATTAAGAGATATTCTTGGTGCATGAGATGTGGTC 1965
 Db 510 ThrSerAlaLeuAspSerGlnSerGlnIleArgGlnAlaLeuAlaArgLeuThr 529
 QY 1966 AAACACAGAACTTCATTTTTCATTCACACAGATTTGCAACAGTGGTTCAGATGAA 2025
 Db 530 ArgLysAspGlnThrThrIleMetIleAlaHisArgLeuSerThrValThrAlaAlaAspAsn 549
 QY 2026 ATCATGTTCTTGATCAGGAGTGAAGTGAAGCCGACGTTGATCCACCAAGTTCCTGCT 2085
 Db 550 IleValIleMetGluGlyGlyGlnValAlaGluGlnGlyPro---GlnGlyArgLeuLeu 568
 QY 2086 AACCTCATGATCTATTCAGAAATGGGATACACAG 2124
 Db 569 SerGlnAspGlyValTyrArgArgLeuTyrGlnLeuGln 581
 RESULT 13
 HLYB_ACTAC ID HLYB_ACTAC STANDARD: PRT: 707 AA.
 AC P23702;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukotoxin secretion ATP-binding protein.
 GN LKTB OR LTB.
 OS Actinobacillus actinomycetemcomitans (Haemophilus
 OS actinomycetemcomitans).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
 OC Actinobacillus.
 NCBI_TaxID=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90384842; PubMed=2402457;
 RA Guthmiller J.M., Kolodrubetz D., Cagle M.P., Kraig E.,
 "Sequence of the lktB gene from Actinobacillus

RT actinomycetemcomitans.";
 RL Nucleic Acids Res. 18:5291-5291 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92072446; PubMed=1961107;
 RA Lally E.T., Golub E.E., Kleba I.R., Taichman N.S., Decker S.,
 Berthold P., Gibson C.W., Demuth D.R., Rosenbloom J.;
 RT "Structure and function of the B and D genes of the Actinobacillus
 RT actinomycetemcomitans leukotoxin complex.";
 RL Microb. Pathog. 11:111-121 (1991).
 CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN (HEMOLYSIN).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. HLYB SUBFAMILY.
 CC -1- SIMILARITY: BEHONGS TO THE ABC TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X53955; CA37906.1; --
 DR PIR, S12601; S12601.
 DR PIR, A61378; A61378.
 DR MEROPS, C39. UNM; --
 DR InterPro: IPR003593; AAA ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001140; ABCtransporter.
 DR InterPro: IPR005074; Peptidase_C39.
 DR Pfam: PF00005; ABC_tran; 1.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF03412; Peptidase_C39; 1.
 DR ProDom: PD000006; ABC_transporter; 1.
 DR SMART, SM00382; AAA, 1.
 DR PROSITE, PS00211; ABC_TRANSPORTER; 1.
 DR Hemolysis; Transport; ATP-binding; Transmembrane.
 FT NP BIND 502 509 ATP (BY SIMILARITY).
 SO SEQUENCE 707 AA; 79578 MW; CA564EACF46DF4FB CRC64;
 Alignment Scores:
 Pred. No.: 4.27e-37 Length: 707
 Score: 651.50 Matches: 173
 Percent Similarity: 52.13% Conservative: 121
 Best Local Similarity: 30.67% Mismatches: 233
 Query Match: 15.31% Indels: 37
 DB: 1 Gaps: 13
 AF133659 (1-2345) x HLYB_ACTAC (1-707)
 QY 463 ATTAGGTTCCCTTCATGTTAAATATGCTGTAGACAGACCTCAACAGATGCGGA--- 519
 Db 169 LeuIleThrProLeuPhePheGlnValAlaMetAspLysValIleValHisArgGlyPhe 188
 QY 520 AACATGCTGAACCTGATGATGATGACCAAAATACAGTTCGAACCATGAGCAAGTTCG 579
 Db 189 SerThrLeuAsn-----ValIleThrValAlaLeuAlaIleVal 201
 QY 580 ATTAGGCTATGCTGATACAGAGCTGAGCTGCTTTTTCATCAAGATTCGAATGCAAGTA 639
 Db 202 ValLeuPheGluIleIleLeuGlyGly-----LeuArgThrTyrVal 215
 QY 640 TTGGGCAAGTAGCCCAAGATTCATCCAGATAGCAAAATATGCTTCCTCATCTT 699
 Db 216 PheAlaHisSerThrSerArgGlyAspValGlnLeuGlnAlaArgLeuPheArgHisLeu 235
 QY 700 CACAACTGATGATGGTGGTTTCACCTGAGACAGACAGACGGAGCTTTATCTAAGGCTATT 759
 Db 236 LeuAlaLeuProIleSerTyrPheGluAlaArgArgValGlyAspThrValAlaArgVal 255
 QY 760 GACAGAGCAACAGGGGTATCATGTTTGTCTGAGTCTTGATTTAATCTTCTCC 819
 Db 256 ArgGluLeuAspGlnIleArgAsnPheLeuThrGlyGlnAlaLeuThrSerIleLeuAsp 275

QY	820	ATCATGTTGAAGTACATCTTGTCAGAGTGGTATTTTGTATTAACAATGGGAGCCAGTTT	879
Db	276	LeuLeuPheSerPheIlePhePheAlaValMetTrpTyr-----SerProIysLeu	293
QY	880	GCTTGGTAACTTGGAAACACTTGGTACATACAAGCATTCACAGTTGACAGTCAACA--	936
Db	294	ThrLeuValValLeuGlySerLeuProCysTrpValIleTrpSerValPheIleSerPro	313
QY	937	-----CGGTGGAACCTAGATTGAAATAGAAATGAAACAAGCAGATTAATGATGACAGT	990
Db	314	IleLeuArgGArgLeuAspAspLysPheAlaArgAsn--AlaAspAsnGln-----	330
QY	991	AATGCTGCTATAGACTCACTGCTGAAATTATGAACCTGGAAGTATTTTAATATGAAGA	105
Db	331	SerPheLeuValGluSerValThrAlaIleAsnThrIleLysSalMetAlaIleSerPro	350
QY	1051	TATGAAACAAGACAGATATGATGATGATTTTGTGAAGACGATGAGACTGCTTCATTGAAAGT	111
Db	351	GlnMetThrAsnIleTrpAspLysGlnLeuAlaSerTyrAlaAlaValSerPheLysVal	370
QY	1111	ACCTTACTCTGGCTAGTGAACCTTGGTCAAGAGTCATATTTCAAGTGGGTTAAAC	117
Db	371	Thr-----ValLeuAlaThrIleGlyGlnGlnGlyIleGlnLeuIle---GlnLys	386
QY	1171	GCTAATAAGTGT-----CTGCCACGTAGGGAATTTGGCAGGTACCTTACT	121
Db	387	AlaValMetValIleAsnLeuTrpLeuGlyAlaHisLeuValIleSerGlyAspLeuSer	406
QY	1219	GTGTGAGATCTAGTATATGATGTAATGGAAGCTGCTTTTACAGCTTCATTACCCCTGAAC	127
Db	407	IleGlyGlnMetIleIleAlaPheAsnMetLeuAlaGlyGlnIleIleSerProValIleArg	426
QY	1279	CTGGGAACGTATATATAGAGACTTACACAAGCACTCATGATATGAAACACCTGTTTACT	133
Db	427	LeuAlaGlnIleTrpGlnAspPheGlnGlnValGlyIleSerValThrArgLeuGlyAsp	446
QY	1339	CTACTCAAGGTAGACACCCCAATTAAGACAAAGATGATGCATCTCCC--CTTCAGATC	139
Db	447	ValLeuAsnSerProTrpGlnAsnAsnThrAlaSerValSerLeuProGlnIleGln--	465
QY	1396	ACACCAACAGACAGCTACCGTGGCTTGATATGATATGCAATTGGAATTC--ATTGAGGGC	145
Db	466	-----GlyGluIleSerPheArgAsnIleLysPheArgTyrLysProAspSer	481
QY	1453	CAGAAAGCTCTTAGTGTGAATTCCTTTGAAGTCCCTCAGAGAAAGAAAGTGGCATTTGA	151
Db	482	ProMetIleLeuAsnAsnIleAsnLeuAspLysIleSerGlnGlyValIleGlyIleVal	501
QY	1513	GGAGGTAGTGGGTACAGGAAAGACAAATAGTAGGCTATTAATTTGGTTCTTATGAGCCT	157
Db	502	GlyArgSerGlySerGlyLysSerThrLeuThrLysLeuIleGlnArgPheTyrIlePro	521
QY	1573	CAAAAGGTACACTTATCTTGCTGCTGTCGTCAAAAATATACAAATGTGAGCCCTGGAACCTT	163
Db	522	GlnGlnGlnGlyValIleLeuIleAspGlyHisAspLeuAlaLeuAlaAspProAsnTrpLeu	541
QY	1633	CGAGAGGCAAGTGGAGTGGTACTCAGAGATGCTGCTCTTCCATTAATACTATTTATAC	169
Db	542	ArgArgGlnValIleGlyValValLeuGlnHisAsnValLeuLeuAsnArgSerIleArgGlu	561
QY	1693	AACCTTATATGAAAACATCAGTGCCTTCACTTGAGGAAGTGTATGACGTGCGCAAAATTA	175
Db	562	AsnIleAlaLeuThrAsnProGlyMetProMetGluLysValIleAlaIleAlaLysLeu	581
QY	1753	GCTGCACTTCAATATGCAATTTCTTGGAAATGCCACATGATATGACACCCAAATGAGGGA	181
Db	582	AlaGlyAlaHisAspPheIleSerGlyLeuArgGlnGlyTyrAsnThrValValGlyGlu	601
QY	1813	CGAGGACCTCAAGCTTTGACAGAGAGAAAGCAAAAGATATACCATTTGCAAGACCATTTTG	187
Db	602	GlnGlyAlaGlyLeuSerGlyGlyGlnArgGlnArgIleAlaIleAlaArgAlaLeuVal	621

QY 1873 AAGAGCCCCCAGCAATCTATGTAAGAGTAACTGATTCATTCGATTCGTA 1933
 Db 622 AaenafProargileuleuileheapgluIatnrSeAlaIeuhspyrGluSerIu 641
 QY 1933 GAGACTATTCGTGGTCGATGAAGAGATGTGTCMAACACAGAACTTATTTTCATTCGA 1992
 Db 642 AauiIleIeMethIaenMethIstlysIleCyGlnAmaargThrValIeuiIeIeMa 661
 QY 1993 CACGATTTGTCACACGTCGTGTATGTCAGATGAATATATTCCTTGATCAGGTAAAGTA 2055
 Db 662 HisArgIeuserThrValIysAaenAlaAaPaArgIleIeIeValMeAAspIsglyIuIe 681
 QY 2053 GCCGACGGTGGTACCACCAATGTTGTCCTTGTCAACCCATAGATATCATTCAGAAAGT 2112
 Db 662 IIdgIuIngIlyshHisgIuIeuleuIysAspGluIyglyIeutyIerIyIeu 701
 QY 2113 TGGCATACACAG 2124
 Db 702 HisgIuIeugIn 705
 RESULT 14
 MSBA_ECOLI STANDARD; PRT; 582 AA.
 ID P2729;
 AC 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable transport ATP-binding protein msba.
 GN MSBA OR B0914 OR Z1260 OR ECS0997.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93172962; PubMed=8094880;
 RA Katow M.L., Georgopoulos C.P.;
 RT "The essential Escherichia coli msba gene, a multicopy suppressor of
 null mutations in the hcrB gene, is related to the universally
 conserved family of ATP-dependent translocators.";
 RL Mol. Microbiol. 7:69-79(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 Kikura S., Kitegawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Moti H., Mochimaru K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
 Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=1074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grotbeck E.J., Davis N.W., Lim A., Dimajanta E.T., Potamoculis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.",
 RL Nature 409:529-533(2001).
 RN
 RC
 RP STRAIN=0157:H7 / RIND 0509952;
 RX MEDLINE=21156231, PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Kida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 RA Kitaura S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.",
 RL DNA Res. 8:11-22(2001).
 [6]
 RP CHARACTERIZATION.
 RX MEDLINE=98241619, PubMed=9575204;
 RA Zhou Z., White K.A., Polissi A., Georgopoulos C., Raetz C.R.,
 RT "Function of *Escherichia coli* MsbA, an essential ABC family
 RT transporter, in lipid A and phospholipid biosynthesis.",
 RL J. Biol. Chem. 273:12466-12475(1998).
 [7]
 RN
 RP X-RAY CRYSTALLOGRAPHY (4.5 ANGSTROMS).
 RX MEDLINE=2143170, PubMed=11546864;
 RA Chang G., Roth C.B.,
 RT "Structure of MsbA from *E. coli*: a homolog of the multidrug resistance
 RT ATP binding cassette (ABC) transporters.",
 RL Science 293:1793-1800(2001).
 CC -1- FUNCTION: INVOLVED IN THE SECRETION OF LIPID A AND POSSIBLY ALSO
 CC GLYCEROPHOSPHOLIPID TRANSPORT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; 211796; CAA7839.1; -.
 DR EMBL; AE000193; AAC74000.1; -.
 DR EMBL; D90729; BA35658.1; -.
 DR EMBL; D90730; BA35660.1; -.
 DR EMBL; AE005281; AAG55399.1; -.
 DR EMBL; AP002553; BAB34420.1; -.
 DR PIR; S21588; S21588.
 DR PIR; S27998; S27998.
 DR PDB; 1J50; 1J50-01.
 DR EcGene; EG10613; mbaA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transport.
 DR InterPro; IPR001140; ABCtransport.
 DR Pfam; PF00005; ABC_tran.1.
 DR Pfam; PF00664; ABC_membrane.1.
 DR SMART; SM00382; AAA.1.
 DR PROSITE; PS00211; ABC_TRANSPORTER.1.
 KW ATP-binding; Transport; Inner membrane; Transmembrane; 3D-structure;
 KW Complete proteome.
 FT DOMAIN 1 21 CYTOPLASMIC.
 FT TRANSMEM 22 52
 FT DOMAIN 53 64 EXTRACELLULAR.
 FT TRANSMEM 65 96
 FT DOMAIN 97 139 CYTOPLASMIC, INTRACELLULAR DOMAIN 1.
 FT TRANSMEM 140 164
 FT DOMAIN 165 167 EXTRACELLULAR.
 FT TRANSMEM 168 192
 FT DOMAIN 193 252 CYTOPLASMIC, INTRACELLULAR DOMAIN 2.
 FT TRANSMEM 253 272
 FT DOMAIN 273 280 EXTRACELLULAR.

FT TRANSMEM 281 301
 FT DOMAIN 302 582 CYTOPLASMIC, INTRACELLULAR DOMAIN 3.
 FT NP_BIND 376 383 ATP.
 SQ SEQUENCE 582 AA; 64460 MW; 3D52A3482174A7FE CRC64;
 Alignment Scores:
 Pred. No: 4.51e-37 Length: 582
 Score: 651.00 Matches: 190
 Percent Similarity: 49.19% Conservative: 115
 Best Local Similarity: 30.65% Mismatches: 259
 Query Match: 15.30% Indels: 56
 Gaps: 17
 AF133659 (1-2345) x MSBA_ECOLI (1-582)
 QY 310 CTCACACAGACCCAAAGAGGTTAAAGATGTGTACTCGAATAATCATAAAGCA 369
 DB 1 Methisambp-----LysAspLeuSerThrPheArg 15
 QY 370 ATGCTTCTTATGTGTGGCCCAAGACGCGCAT-----CTACGACTAGA 417
 DB 16 Leu-----TTPProThrIleAlaProPheLysAlaGlyLeuIleValIleGly 31
 QY 418 GTTGCCATTGCTGGGATTTTGGGTGGCAAGCCATGAATATTGGCTTCCTTC 477
 DB 32 ValAlaLeuIleLeuAsnAlaLaserAspThrPheMetLeuSerLeuLysProLeu 51
 QY 478 ATGTTTAAATATGTGTATGACAGCCTCAACAGATGCGGAAACATGCTGAACCTGAGT 537
 DB 52 Leu-----AspAsp-----GlyPheGly 57
 QY 538 GATGACCAATAATGATGATGACCAACCATGACAGCAGTTGATGCTATGCTATGATCA 597
 DB 58 LysThrAspArgSerValIleValIleThrPheProLeuValIleGlyLeuMetIleLeu 77
 QY 598 AGAGTGGAGCTGCTTTTAAACGAA---GTGGAATAGCATTTTGGCACTAGCC 654
 DB 78 ArgGlyIleThrSerIleValSerSerIleCysIleSerIleValSerIleValIleVal 97
 QY 655 CAGAAATCAATCCAGAGATGACCAAAATATGCTTTCATCTTCAACACCTGATCTGT 714
 DB 98 MetThrMetAlaArgAlaGluLeu-----PheGlyIleMetMetGlyMetProVal 113
 QY 715 GGTTCACCTGACAGACGACGAGCCTTATCTAAGCTATGACAGGAACAGG 774
 DB 114 SerPhePheAspLysGlnSerThrGlyThrLeu-----LeuSerArgIleThrTyr 130
 QY 775 GGTATCAGTTTGTCTGAGTCTTGGTATTTATCTTCTCCCATCATCTTGAAGTG 834
 DB 131 AspSerGlnGlnValAlaSerSerSerSerGlyAlaLeuIleThrValAlaArgGly 150
 QY 835 ATGCTGTCACTGCT-----GTTTGTATTAACAATGCGGCGCCACTTGGCT 882
 DB 151 AlaSerIleIleGlyLeuPheIleMetMetPheIleTyrIleSerThrGlnLeuSerIleIle 170
 QY 883 TTGTAAACCTTGAACACTTGTATACATACACAGCATTCACAGTGCACGCGTGG 942
 DB 171 LeuIleValLeuAlaProIle-----ValSerIleAlaIleArgValVal 185
 QY 943 AGAATGATTTAAGATTAAGAAATGAACAAACAGATTAATGACAGT-----AAT 993
 DB 186 SerLysArgPheArgSerIleSerLysAsnMetGlnSerThrMetGlyGlnValThrThr 205
 QY 994 GCTGCTATAGACTACGCTGTAATTAATGAAGAGTATTAATTAATGAAGATAT 1053
 DB 206 SerAlaGlnGlnMetLeuLysGlyIleValIleValIlePheGlyGlnGlnVal 225
 QY 1054 GAACACACAGATATGATGATTTTGAAGACGTATGACGCTTCAATGAAGAAAGTACC 1113
 DB 226 GluThrLysArgPheAspLysValSerAsnArgMetLysGlnGlyMetLysMetVal 245
 QY 1114 TCTACTGCTATGCTGAAGCTTGTGCAAACTGCTATTTTCAGTGTGCTTTAAACAGCT 1173
 DB 1114 TCTACTGCTATGCTGAAGCTTGTGCAAACTGCTATTTTCAGTGTGCTTTAAACAGCT 1173

```

Db 246 SerAlaSerSerIleSerAspProIleIleGlnLeuIleAlaSerLeuAlaLeuAlaPhe 265
QY 1174 ATAAATGTCCTCCGACGTCAGGAATGTGGACAGTACCTTACTGTTGAGATCTAGTA 1233
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 266 ValLeuTyrAlaIleAspPheProSerValIleAspSerLeuThrAlaGlyThrIleThr 285
QY 1234 ATGTGAATGACGCTCTTTTTCAGCTTTCATTACCCCTGAACTTTCTGGGAACTGTATAT 1293
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 286 ValValPheSerSerMetIleAlaLeuMetArgProLeuTyrSerSerLeuThrAsnValAsn 305
QY 1294 AGAAGACTAGACAGACCATGATGATATGAAACCTTGTCTTACTTACTTACGAGTGAGAC 1353
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 306 AlaGlnPheGlnArgGlyMetAlaAlaCysGlnThrLeuPheThrIleLeuAspSerGln 325
QY 1354 ACCCAATTAAGAACAAGTATGATGGCATCTCCCTTCAGATCACACACAGACGCTACC 1413
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 326 ---GlnGluTyrAspGlu-----GlyTyrArgValIleGluAlaGlyAlaThrGlyAsp 341
QY 1414 GTGGCTTTGATTAATGTGCATTTTGAATATACATTAGAGGCCAGAAAGTC-----CTTAGT 1467
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 ValGluPheArgAsnValIleThrPheThrTyr---ProGlyArgAspValProAlaLeuArg 360
QY 1468 GGAATATCCTTTGAAGTCCCTCCGACAGAAAGAGTGGCCATGTAAGAGTATGAGGTCA 1527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AsnIleAsnLeuTyrSileProAlaGlyLysThrValAlaLeuValGlyArgSerGlySer 380
QY 1528 GGGAAAGACAATATGATGAGGCTATTTTCGCTTCTATGAGCCTCAAAAAGGTAGCATT 1587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 GlyLysSerThrIleAlaSerLeuIleThrArgPheTyrAspIleAspGluGluIle 400
QY 1588 TATCTTCTGCTCAAAATATATCAAGATGTAGCCTGGAAGCCTCGAGGGCAGTGGGA 1647
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 401 LeuMetAspGlyHisAspLeuArgGlyTyrThrLeuAlaSerLeuArgAsnGlnValAla 420
QY 1648 GTGTACTACTAGAGATGCTGTCCTCTTCATATATCTATTTATTCACACCTCTTATAGGA 1707
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 421 LeuValSerGlnAsnValHisLeuPheAsnAspThrValAlaAsnAlaIleAlaTyrAla 440
QY 1708 AACATCAGTGTCT--TCACCTGAGGAAGTATGATGACATGGCAAAATTAGTGACTTCAT 1764
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 441 ArgThrGluGlnTyrSerArgGluGlnIleGluGlnAlaAlaArgMetAlaTyrAlaMet 460
QY 1765 GATGCAATTCTTGAATGCCATGATATATGACACCAAGTAGGGGACAGAGACTGCAAG 1824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 AspPheIleAsnLysMetAspAsnGlyLeuAspThrValIleGlyGluAsnGlyValLeu 480
QY 1825 CTTTCAGAGAGAGAAAGCAAGAGTAGCAATTCAGAGAGCATTTCAGAGAGACCCCA 1884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 LeuSerGlyGlnArgGlnArgIleAlaIleAlaArgAlaLeuLeuArgAspSerPro 500
QY 1885 GTCATACTATATGTAAGGCTACTTCACTGTTAGTTCGATTCGATTAAGAGACTATCTT 1944
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 501 IleLeuIleLeuAspGluAlaThrSerAlaLeuAspThrGluSerGluAlaIleGln 520
QY 1945 GGTGCCATGAGAGATGTGTCAACACAGAACTTCTATTTTCATTGACACAGATTGTCA 2004
    ::::: ||||| ||||| ::::: ||||| ||||| ::::: ||||| |||||
Db 521 AlaAlaLeuAspGluLeuGlnLysAsnArgThrSerLeuValIleAlaHisArgLeuSer 540
QY 2005 ACAAGTGTATGAGATGAATATATGTCTGTGATCAGGGTAAAGTAAAGTAAAGTGT 2064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 ThrIleGluTyrAlaAspGluIleValIleValIleGluAspGlyValIleValIleGluArgGly 560
QY 2065 ACCCAATGATGTTGCTTGTGAACCTTCATATCTATTCAGAAATAGTGGCATATACAG 2124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 561 ThrHisAsnAspLeuLeuGluHis--ArgGlyValTyrIleGlnLeuHisLysMetGln 579

```

```

DE Leukotoxin secretion ATP-binding protein.
GN LKTB.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=28165;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=9329320; PubMed=8478098;
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT "Molecular characterization of a leukotoxin gene from a Pasteurella
RT haemolytica-like organism, encoding a new member of the RTX toxin
RT family."
RL Infect. Immun. 61:2089-2095(1993).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L12148; AAA16445.1; -
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001140; ABCtransportTM.
DR InterPro; IPR005074; Peptidase_C39.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF03412; Peptidase_C39; 1.
DR ProDom; PD000006; ABC_transport; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR HEMOLYSIS; Transport; ATP-binding; Transmembrane.
FT TRANSMEM 159 176 POTENTIAL.
FT TRANSMEM 189 212 POTENTIAL.
FT TRANSMEM 270 288 POTENTIAL.
FT TRANSMEM 295 316 POTENTIAL.
FT TRANSMEM 388 404 POTENTIAL.
FT TRANSMEM 411 428 POTENTIAL.
FT NP BIND 503 510 ATP (BY SIMILARITY).
SQ SEQUENCE 708 AA; 79772 MW; 7326D2141B5090A CRC64;

Alignment Scores:
Pred. No.: 7,41e-37 Length: 708
Score: 648.00 Matches: 202
Percent Similarity: 47.14% Conservative: 144
Best Local Similarity: 27.52% Mismatches: 278
Query Match: 15.23% Indels: 110
DB: 1 Gaps: 24

AF133659 (1-2345) x HLYB_PASSP (1-708)
QY 31 TGGGCGCGCGCGGCTGCTTTGAA---AGCGCGGACCTCCGAGATTCTGATCCG 87
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47 TrpLeuLeuAlaIleLysSerLeuGluLysAlaLysHis-----IleLys 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 88 CTTTAGTCTCTGTAGCGGCTCAGGTCCGAGTGAAGGCATCACTCGGCGCTT 147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 LysGluIleSer-----ArgLeuHisLeuVal----- 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 148 GGAACGCTCGAGCTTACAGATTCCAGAGTCATTAATAAGTATCAGTCGACGAGATTG 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 -----AsnLeuProAlaLeuValTrpGlnAspAsn 81
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 208 GGAAGAAGCAATTCAGACAGTTCTTA----- 234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 GlyLys-----HisPheLeuLeuValLysIleAspThrAspLysLysArgTyr 97

```

QY 235 -----GATGTCGAAAGGCTCTCCAGTATGCGCCACTGATAGAA 273
Db 98 LeuThrTyrAsnLeuGluGlnAspAlaProlysIleLeuSer-----GlnGluGln 114
QY 274 AAGAGACATGTTGGCATGTCATGCAGAGAGAGACTCCACACAGACCCAAAGAGGG 333
Db 115 PheGluSerCysTyrGlnGlyLys----- 122
QY 334 TTTAAAGATGTGATACGCGAAATCATTAAGCAATGCTTCTC-----TATGTG 384
Db 123 IleIleLeuValThrSerArgAlaSerIleValGlyGlnLeuAlaLysPheAspPheThr 142
QY 385 TGG-----CCCAAGACAGCCGAGATCTACGAGCTAGAGTTGCCATTTCG 429
Db 143 ThrPheIleProAlaValIleLysTyrThrGlySilePheLeuGlnThrLeuLeuValSer 162
QY 430 CTGGGATTTTGGTGGTGCAGAAAGCCATGAATATGTGTGCTTCCTTCATGTTAAATAT 489
Db 163 Ile---PheLeu-----GlnIlePheAlaLeuIleThrProLeuPheGlnVal 178
QY 490 GCTGAGACAGCCTCAACCAAGATGTCGGA---AACATGCTGAACCTGAGTATGACCA 546
Db 179 ValMetAspLysValLeuValHisArgGlyPheSerThrLeuAsn----- 193
QY 547 AATACAGTTGCACCATGCGACACAGCAGTTCTGATGCTATGCTATCAAGACTGA 606
Db 194 -----IleIleThrValAlaLeuAlaIleValIleIlePheGlnIleValLeuSerGly 211
QY 607 GCTGCTTTTAAAGAAAGTTCGAAATGCAATGATTTGGCAGAGTACCAGAAATTCATC 666
Db 212 -----LeuArgThrTyrIlePheAlaHisSerThrSerArgIleAsp 225
QY 667 CGAAGATAGCCAAAGATGCTTCTTCATCTTCACAACTGAGATCGGTTTCACTCG 726
Db 226 ValGluLeuGlyAlaArgLeuPheArgHisLeuLeuAlaLeuProIleSerTyrPheGln 245
QY 727 AGCAGACAGACGAGAGCTTATCTTAAGGCTATGACAGAAACAAAGGATTCAGTTT 786
Db 246 AsnAspArgValGlyAspThrValAlaArgValArgGluLeuAspGlnIleArgAsnPhe 265
QY 787 GTCCTGAGTGTCTTGATTAATCTTCTCCATCATGTTGAAGTATGCTTCACT 846
Db 266 LeuThrGlyGlnAlaLeuThrSerValLeuAspLeuPheSerPheIlePhePheAla 285
QY 847 GGTGTTTGTATTAACAATGCGTCCGAGTGTGCTTGTGTAACCTTGGAACACTTGGT 906
Db 286 ValMetTyrTyr-----SerProLysLeuThrLeuValIleLeuGlySerLeuPro 303
QY 907 ACATACACAGCATTCACAGTTGCATGCACACGTCGAGAACTAGATTGAATAGAA--- 963
Db 304 CysTyrIleLeuThrSerIlePheIleSerProIleLeuAspArg---ArgLeuAspAsp 322
QY 964 -----ATGAACAAAGCAGATATGATGAGGTAATGCTGTATAGACTCACTGCTGAAT 1017
Db 323 LysPheAlaArgGlyAlaAspAsnGlnAla-----PheLeuValGlnSerValThrAla 340
QY 1018 TATGAACCTGTAAGTATTTTAATGAAGATATGAAGCACAAGATATATGATTT 1077
Db 341 IleAsnMetIleLysAlaMetAlaValSerProGlnMetThrAspThrTyrAspLysGln 360
QY 1078 TTGAAGAGTATGAGTACCTCATTTGAAAGTACCTTACTGCTAGTATGTAACCTT 1137
Db 361 LeuAlaSerTyrValSerSerSerPheArgValThr-----ValLeuAlaThrIle 377
QY 1138 GGTCAAGGTGATTTTTCAGTGTGCGTTTACAGCTATATATGCTG-----CTCGCC 1188
Db 378 GlyGlnGlnGlyValGlnLeuIleGlnLysThrValMetValIleAsnLeuThrLeuGly 397
QY 1189 AGTCGGAATTTGCGAGGTACCTTACTGTTGAGATCTGTGAATGATGATGATG 1248
Db 398 AlaHisLeuValIleSerGlyAspLeuSerIleGlyGlnLeuIleAlaPheAsnMetLeu 417

QY 1249 CTTTTCAGCTTTTCATTACCCCTGAACTTTCTGGAACTGTATATGAGACTAGACAA 1308
Db 418 SerGlyGlnValIleAlaProValIleArgLeuAlaGlnLeuThrGlnAspPheThrGln 437
QY 1309 GCACCTCATAGTATGAACACCTTTGTTACTCTACTCAAGGTAGACACC---CAATTTAA 1365
Db 438 ValGlyIleSerValThrArgLeuGlyAspValLeuAsnSerProThrGlnGlnTyrGln 457
QY 1366 GACAAAGTATGATGCTCCCTTCAGATCACACACAGACAGCTTACCGCTTGTAT 1425
Db 458 GlyLysLeuSerLeuProGlnIleGln-----GlyAspIleAlaPheLys 472
QY 1426 AATGTGATTTTGAATAC---ATTGAGGCCAGAAAGTCTTATGGAATATCTTTGAA 1482
Db 473 AsnIleArgPheArgTyrLysProAspAlaProThrIleLeuAsnValAsnLeuGln 492
QY 1483 GTCCCTGCAGAAAGAAAGTGCATTTGTAGAGTGTGGGTGAGGAAAGACATA 1542
Db 493 IleLysLysGlyGluValIleGlyIleValGlyArgSerGlySerGlyLysThrLeu 512
QY 1543 GTGAGGCTATATTGCTGCTTCTTATGAGCTTCAAAGGGTATGATTTTCTGTGTCAA 1602
Db 513 ThrLysLeuLeuGlnArgPheTyrIleProGlnAsnGlyGlnValLeuIleAspGlyHis 532
QY 1603 AATATACAGATGTGAGCTGGAAGCCTTCGAGGSCAGTGGAGTGTACTTCAGAT 1662
Db 533 AspLeuAlaLeuAlaAspProAsnThrLeuArgGlnIleGlyValIleGlnAsp 552
QY 1663 GCTGTCTCTTCCATATACATTTATTTATCAACTCTTATATGAAACATCAGCTTCA 1722
Db 553 AsnValIleLeuAsnAspArgSerIleArgGlnAsnIleAlaLeuSerGluProGlyMetSer 572
QY 1723 CCTGAGGAAGTATGACAGTGCAGAAATTTAGCTGACTTCATGCAATCTTCGAT 1782
Db 573 MetGluArgValIleTyrAlaAlaLysLeuAlaGlyAlaHisAspPheIleSerAspVal 592
QY 1783 CCACATGATATGACACCCCAAGTAGGGGAACGAGACTCAAGCTTTCAGAGAGAAAG 1842
Db 593 ArgGlnGlyTyrAsnThrIleValGlyGlnGlnIleValGlyLeuSerGlyGlyArg 612
QY 1843 CAAAGATGACATTTGCAAGAGCATTGTAAGAACCCCCAGTCATATCTATGATGAA 1902
Db 613 GlnArgIleAlaIleAlaArgAlaLeuValAsnAsnProLysIleLeuIlePheAspGln 632
QY 1903 GCTACTTATGCTATGATTCATTAAGAGACTATTCTTGTGTCATGAGATGAT 1962
Db 633 AlaThrSerAlaLeuAspTyrGlnSerGlnHisIleIleMetGlnAsnMetGlnLysIle 652
QY 1963 GTCAACACAGAACTTCTATTTTCATTTGCACACAGATTTGCAACAGTGTGATGCAT 2022
Db 653 CysGlnGlyArgThrValIleLeuIleAlaHisArgLeuSerThrValLysAsnAlaAsp 672
QY 2023 GAAATCATTTGCTTGATCAGAGGTAAAGGTAGCCAGTGTATCCCACTGTTGCTT 2082
Db 673 ArgIleIleValMetGluLysGlyGlnIleValAlaGlnGlnGlyLysHisAsnGlnLeu 692
QY 2083 GCTAACCTCATAGTATCTTCAAGAAATGCGCATACAG 2124
Db 693 GlnAsnAsnGlyLeuTyrSerTyrLeuHisGlnLeuGln 706

Search completed: February 15, 2003, 05:54:54
Job time : 48 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 05:51:35 ; Search time 46 Seconds
(without alignments)
9801.531 Million cell updates/sec

Title: AF133659
Perfect score: 4254
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAATCATACATT 2345

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p model -DBV=xih
-Q=/cgr2.1/USFTO_spool/AF133659/runat_10022003_155441_29535/app_query.fasta_1.2503
-DB=PIR_73 -QMT=fastran -SUFFIX=xpr -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blomum2 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=AF133659 @CGN 1.1 60 @runat_10022003_155441_29535 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_NMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DBV TIMEOUT=120
-MAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1657.5	39.0	677	2	T10657
2	1620.5	38.1	680	2	T10656
3	1538.5	36.2	693	2	T39154
4	1516.5	35.6	690	2	S54211
5	1454.5	34.2	631	2	AH0315
6	1446.5	34.0	609	2	G71731
7	1400	32.9	628	2	AF3438
8	1375	32.3	629	2	AG2707
9	1375	32.3	629	2	H97489
10	1247	29.3	836	2	JEO248
11	1042	24.5	643	2	C87412
12	970.5	22.8	830	2	S25198
13	923.5	21.7	1025	2	T18376
14	915	21.5	772	2	T26313

15	725	17.0	642	2	F84172	ABC transport prot
16	709	16.7	581	2	E83790	ABC transporter (A
17	708	16.6	597	2	AD2074	ATP-binding protei
18	706.5	16.6	593	2	S75352	ABC-type transport
19	700	16.5	832	2	B87673	ABC transporter, H
20	699.5	16.4	578	2	A89874	hypothetical prote
21	687	16.1	585	2	H69857	ABC transporter (A
22	678.5	15.9	602	2	A75590	ABC transporter, A
23	675.5	15.9	594	2	S72638	hypothetical ABC e
24	674.5	15.9	604	2	F63802	ABC transporter (A
25	673.5	15.8	600	2	AD3138	hypothetical prote
26	673.5	15.8	600	2	H98149	ATP-binding transp
27	673.5	15.8	863	2	C90482	ABC transporter, A
28	672.5	15.8	546	2	AG1886	ATP-binding protei
29	672.5	15.8	707	1	LEECB	hemolysin secretio
30	671.5	15.8	611	2	AH3540	ABC transporter AT
31	671.5	15.8	707	2	S10057	ABC-type transport
32	670.5	15.8	584	2	E83767	ABC transporter (A
33	666.5	15.7	599	2	D98267	ABC transporter AT
34	666.5	15.7	599	2	AD3017	hypothetical prote
35	664	15.6	600	2	AI2454	ATP-binding protei
36	662	15.6	595	2	T31077	probable ABC-trans
37	662	15.6	650	2	T17482	ABC-type transport
38	658	15.5	708	2	C30169	leukotoxin express
39	654	15.4	582	2	AF0614	probable transport
40	653	15.4	893	2	AH2007	toxin secretion AB
41	652	15.3	607	2	A95911	hypothetical prote
42	651.5	15.3	707	2	A61378	leukotoxin express
43	651	15.3	582	2	S27998	probable ABC-type
44	651	15.3	582	2	C85617	ATP-binding transp
45	651	15.3	582	2	E90753	ATP-binding transp

ALIGNMENTS

RESULT 1
T10657
probable ABC-type transport protein T5F17.80 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C/Accession: T10657
R/Bevan, M.; Demard, N.; Quail, M.; Harrie, B.; Rajandream, M.A.; Barrell, B.G.; Bancr
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16533
A/Accession: T10657
A/Molecule type: DNA
A/Residues: 1-677 <BEV>
A/Cross-references: EMBL:AL049917; GSPDB:GN00062; ATSP:T5F17.80
A/Experimental source: cultivar Columbia; BAC clone T5F17
C/Genetics:
A/Gene: ATSP:T5F17.80
A/Map position: 4
A/Intons: 519/3
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
F/452-646/Domain: ATP-binding cassette homology <ABC>

Alignment Scores:

Pred. No.:	3.52e-112	Length:	677
Score:	1657.50	Matches:	326
Percent Similarity:	74.30%	Conservative:	128
Best Local Similarity:	53.36%	Mismatches:	152
Query Match:	38.96%	Indels:	5
DB:	2	Gaps:	4

AF133659 (1-2345) x T10657 (1-677)

QY	316	ACAGACCCAAAGAGGGTTAAAGATGTTGATCTCGGAAATCATMAAGCAATGCTT	375
DB	68	ThrProandlnspglnthrLysstrAlaserSerLysIleuauGthrIleSer	87
QY	376	TCTTATGTGTGGCCCAAGACAGCCAGATCTACGAGCTGAGTTGGCATTTTGGCTGGGA	435
DB	88	SerTyrIeuThrMetLysAspAenProGlnLeuAArgPheArValIleAlaLeuAla	107

```

QY 436 TTTTGGGTGTCGAAGGCCATGATATGTGCTCCCTCATGTTTAATATGCTGTA 495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 CysleuIleGlyAlaIlysheuIeuIeuValGlnValProPheIeuPheIyIseuIle 127
QY 496 GAGACCTCAACCAAGATGTGGGAAACATGCTGAACCTGAGTATGACCAAAATACAGT 555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 AspleuIeuSerSerIySerSerSerThrIleThrAspSerAsn---ProIyIleuIeu 146
QY 556 GCAACCATGGAACA-----GCAGTTCGATTGGCTATGCTATGCTACAGCTGAGCT 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 AlaAlaPheAlaThrProSerSerValIleuIleGlyIyGlyIleAlaIySerIySer 166
QY 610 GCTTTTAAACGAAGTTGGAATGCAAGTATTTGGACAGTGGCCAGCAATTCGCA 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 167 SerAlaPheAsnGlnIeuIyGlnIyAlaIyPheSerIyValSerIeuIyGlnIyIleIy 186
QY 670 AGAATAGCCAAAATGTCTTCTTCATCTTCAATCTTCAACCACTGGATCTGGGTTTCACTGAGC 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 SerValSerArgIyValIleuSerIyIleuIyIleuIyIleuIyIleuIyIleuIy 206
QY 730 AGACAGAGCGGAGCTTTATCTAAGCTATGACAGAGAAACAGGGATATCATGTTTGC 789
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 ArgGlnIyGlyAlaIleuIyIleuIyIleuIyIleuIyIleuIyIleuIyIleuIy 226
QY 790 CTGAGTGTCTTGGTATTTAATCTTCTTCATCATGTTTGAAGTATGCTGCTGAGT 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 LeuSerAlaMetValPheAsnValValProThrIleIleuGlnIleSerMetValIyIle 246
QY 850 GTTTGTGTTTCAAAATGCGGTGCCAGTGTGCTTGTGTAACCTTGAACACTTGCTGACA 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 IleIeuAlaIyAsnIyPheGlyIyProValPheAlaIleuIleThrSerIeuSerValIyIySer 266
QY 910 TACACGAGATTACAGTGCAGTGCAGGAGTGCAGATGTAATGAATGAATGAC 969
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 TyrIleAlaPheThrIleuValIyAlaIyGlnIyIyIyIyIyIyIyIyIyIyIyIy 286
QY 970 AAAGCAGATTAATGATGACGATGATGCTGCTATGATGATGATGATGATGATGATGATG 1029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 GlnAlaPheAsnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 306
QY 1030 AAGTATTTAATGAAGAATGATGACAGATGATGATGATGATGATGATGATGATGAT 1089
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 LysIyIyIyPheAsnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 326
QY 1090 GAGATGCTTCACTTGAAGAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 GlnAspAlaAlaIleuGlnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 346
QY 1150 ATTTTCAGTGTGCTTTAACAAGCTAATATGCTGCTGCGCAGTCAAGGAAATGTCGACAGT 1209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 IlePheSerThrAlaIleuSerThrSerMetValIleuIyIyIyIyIyIyIyIyIyIyIy 366
QY 1210 ACCCTTACTGTGAGATCTAGTAATGTGTAATGTGTAATGTGTAATGTGTAATGTGTAAT 1269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 GlnMetThrValGlyAspIleuValMetValIyIyIyIyIyIyIyIyIyIyIyIyIy 386
QY 1270 CTGAACCTTCTGGGAACTGATATATAGAGACTAGACAAGCACTCATGATATGACAACC 1329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 LeuIyIyPheIleuGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 406
QY 1330 TTGTTTACTTACTCAAGATGACACCCCAATTAAGCAAAAGTATGATGATGATGATGATGAT 1389
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 LeuPheGlnIleuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 426
QY 1390 ---CAGATCACACACAGACAGTACCGTGGCTTGTGTAATGTGTAATGTGTAATGTGTAAT 1446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 ProIleuValIleuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 446
QY 1447 GAGGCGCAAGAAAGTCTTATAGTGAATATCTTGAAGTCCCTGCGAGGAAAGAAAGTGGCC 1506
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 447 ProGlnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 466

```

```

QY 1507 ATTGTAGAGGTAGTGGGTGACGAGGAAAGCAAAATAGTACGCTATTTATTCCTCTAT 1566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 IleValIySerSerGlyIySerGlyIySerThrIleIeuIyIyIyIyIyIyIyIyIyIyIy 486
QY 1567 GAGCCTCAAAAAGGTAGATTTATCTGCTGCTGCAAAATATATACAAATGTGACCTGGAA 1626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 AspThrAspSerGlyAsnValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 506
QY 1627 AGCCTTGGAGGAGGAGGAGGAGTACCTGAGATGCTGCTCTCTCTCTCTCTCTCTCTCT 1686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 507 SerIeuIySerCysIleGlyIyAlaIyProIyIyIyIyIyIyIyIyIyIyIyIyIyIy 526
QY 1687 TATTACAACCTCTTATATGAAACATGATGCTTCACTGAGGAAAGTATGACAGTGGCA 1746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 PheIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 546
QY 1747 AAATTAGCTGATCTTCAATGATGCAATTTCTTGAATGCCACATGATATGACACCCAACTA 1806
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 ArgArgIyAlaValIleIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 566
QY 1807 GGGGACGAGGACTCAAGCTTTTTCAGAGGAAAGCAAGAGATGCAATTCGCAAGGCC 1866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 567 GlyIyIyIyGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 586
QY 1867 ATTTTGAAGACCCCGCAGTCACTATGATGATGATGATGATGATGATGATGATGATGAT 1926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 587 PheIeuIySerProIyAlaIleIleuIyCysAspIyIyIyIyIyIyIyIyIyIyIyIyIy 606
QY 1927 ACTGAGAGACTATTTCTTGGTGCATGAGATGCTGCAACACAGAACTTCTATTTTC 1986
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 ThrGlnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 626
QY 1987 ATTCACACAGATGTCACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 2046
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 627 IleAlaIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 646
QY 2047 AAGGTACCGCAAGCTGTGATCCACCATGATGTTGCTTCAACCTCATGATATCTATTC 2106
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 647 LysValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 665
QY 2107 GAAATGTGGCATACACAGACGCGGTGCGAG 2119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 666 LysIeuIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 676

```

RESULT 2

```

probable ABC-type transpore protein T5F17.70 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear crease)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #ext_change 21-Jan-2000
C/Accession: T10656
R/By: M. J. Lennard, N. Q. Quail, M. J. Harris, B. Rajandream, M. A. Barrell, B. G. Bancr
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16533
A/Accession: T10656
A/Molecule type: DNA
A/Residues: 1-680 <BEV>
A/Cross-references: EMBL:AL049917; GenBank:U00062; ATSP:T5F17.70
A/Experimental source: cultivar Columbia; BAC clone T5F17
C/Genetic:
A/Map position: 4
A/Insertions: 522/3
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
F/455-649/Domain: ATP-binding cassette homology <ABC>

```

Alignment Scores:

```

Pred. No.: 1,69e-109 Length: 680
Score: 1620.50 Matches: 321
Percent Similarity: 71.32% Conservative: 129
Best Local Similarity: 50.87% Mismatches: 170
Query Match: 38.09% Indels: 11
DB: 2 Gaps: 5

```

AF133659 (1-2345) x T10656 (1-680)

QY 256 GTATGGCCATGATGAAAGAGACATGTTGGATGTCAGAGAGAGACTCCAC 315
 Db 57 ILeThrProMetValAsnAlaArgValMeTherPheSerThrSerAlaProHis 76
 QY 316 ACAGACCCAAAAGGGTTAAAGATTTGATCTCGGAAATCATTAAGCAATGCTT 375
 Db 77 -----ProGluValLeuAsnArgThrSerSerGluAsnIleLeuArgMetIleSer 93
 QY 376 TCTTATGTTGGCCCAAGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 435
 Db 94 SerTrpLeuTrpMetLysAspAsnProLysLeuLysCysPheArgValIleSerAlaPheAla 113
 QY 436 TTTTGGTGGTGGCAAGGCCATGATATGTTGGTCCCTTCATGTTAAATATGCTGTA 495
 Db 114 CysLeuValGlyIleAlaLysPheLeuAsnValGlnIvalProPheLeuPheLysValAlaIle 133
 QY 496 GACAGCCTCAACCAAGATGTCGGGAAACATGCTGAACCTGAGTATGACCAATACAGTT 555
 Db 134 AspTrpLeuSerSerSerSer-----PheValIleSerAsnProTrpLeuVal 149
 QY 556 GCAACCATGAGACA-----GCAGTTCTGATGGCTATGCTATGCTATGATGATGATGATGAT 609
 Db 150 AlaAlaPheAlaTrpProSerSerValLeuIleGlyTrpGlyIleAlaArgSerGlySer 169
 QY 610 GCTTTTAAAGCAAGTTGCAAAATGCAATGATTTGGCAAGTAGCCAGAAATTCATCCGA 669
 Db 170 SerAlaPheAsnGluLeuArgThrSerValPheSerIysValAlaLeuArgThrIleArg 189
 QY 670 AGAATGACCAAAAATGCTCTTCTCCATCTTCAACCTGATGCTGGGTTTTCACCTGAGC 729
 Db 190 ThrIleSerArgLysValLeuSerArgLeuHisAspLeuAspLeuArgLysIleLeuAsn 209
 QY 730 AGACAGAGGGAGGTTTATCTAAAGCTATTGACAGAGAAACAAGGGATACAGTTTGTG 789
 Db 210 ArgAspThrGlyAlaLeuAsnArgIleIleAspArgLysSerAlaIleAsnThrIle 229
 QY 790 CTGAGTCTGTTGATTTAAATCTCTCCCATCATGTTTGAAGATGATGCTGTCAGTGTG 849
 Db 230 LeuSerAlaMetValPheAsnIleMetProThrIleLeuGluIleSerMetValSerCys 249
 QY 850 GTTTTGTATTCAAAATGCGGTGCCAGTTGCTTTGGTAAACCTTGAAACCTTGATGTA 909
 Db 250 IleLeuAlaTrpLysPheGlyAlaValTrpAlaLeuIleThrCysLeuSerValGlySer 269
 QY 910 TACACAGATTCAAGTTGACAGTACACAGGTGGAGAACTAGATTAGATTAAGAAATGAC 969
 Db 270 TyrIleAlaPheThrLeuAlaMetThrGlnTrpArgIleLysIleArgLysAlaMetAsn 289
 QY 970 AAACGAGTATATGATGAGGATATGCTGCTATGATCACTCACTGATGATTTTGAACGTG 1029
 Db 290 GluAlaGluLeuAsnAspAlaSerThrArgAlaIleAspSerLeuIleAsnTrpGlnThrVal 309
 QY 1030 AAGTATTTTAAATGAAGAATATGAAGACAGAGATATGATGATTTTGAAGACCTAT 1089
 Db 310 LysTrpThrPheAsnAsnGluAspTrpGluAlaArgLysTrpAspGlnLeuHisGluAsnTrp 329
 QY 1090 GAGACTGCTTCAATGAAAGTACTTACTGCTGCTATGCTGAACTTTGGTCAAAAGTCT 1149
 Db 330 GluAspAlaIleLeuGlnSerArgLysSerPheAlaLeuLeuAsnPheGlyGlnSerPhe 349
 QY 1150 ATTTTCACTGCGGTTTAAACGATATATGATGCTGCTGCCAGTCAAGGAAATTTGGCAGGT 1209
 Db 350 IlePheSerThrAlaLeuSerThrAlaMetValLeuCysSerGlnGlyIleMetAsnGly 369
 QY 1210 ACCCTTACTTTGGAGATCTAGTATATGATGATGATGATGATGATGATGATGATGAT 1269
 Db 370 GluMetThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeuPro 389
 QY 1270 CTGACACTTTTGGGACTGTATATAGAGACATGACAGACACTCATAGATATGACACACC 1329
 Db 390 LeuTrpThrLeuGlyValValTrpSerAspThrValGlnGlyLeuValAspMetLysSer 409

QY 1330 TTGTTTACTCTACTCAAGGTAGACCCCAATTAAAGACAAAGTATGATGATGATGATGAT 1389
 Db 410 MetPheLysPheLeuGlnGluArgSerAspIleGlyAspLysAspIleAspArgLysLeu 429
 QY 1390 ----CAGATCACACACACAGACAGCTACCGTGGCCCTTTGATATATGTCATTTGAAATACATT 1446
 Db 430 ProProLeuValLeuLysGlyGlySerIleSerPheGluAsnValHisPheSerTrpLeu 449
 QY 1447 GAGGCGCAAGAAAGTCTTATGGAATATCTTGAAGTCCCTGAGAGAAAGAAAGTGGCC 1506
 Db 450 ProGluArgLysIleLeuAspGlyIleSerPheGluValProAlaGlyLysSerValAla 469
 QY 1507 ATTGAGAGAGTATGAGTCAAGGAGAAAGCAATATGTCAGGCTATATTTTGGCTTCTAT 1566
 Db 470 IleValGlySerSerGlySerGlyLysSerThrIleLeuArgMetIlePheArgPhePhe 489
 QY 1567 GAGCCTCAAAAAGGTACATTTATCTTGTGCTGCAAAAATATACAGATGTGAGCCTGAA 1626
 Db 490 AspValAspSerGlyAsnValLysIleAspGlyGlnAspIleLysGluValArgLeuGlu 509
 QY 1627 AGCCTTGAGAGGAGAGGAGTACCTCAGAGATGCTGCTCCCTTCCATTAATCTATT 1686
 Db 510 SerLeuArgSerSerIleGlyValValProGlnAspThrValLeuPheAsnAspThrIle 529
 QY 1687 TATTAACAACCTCTTATATGAAACATCAGTGTCTTCACTGAGAGAGTATGACAGTGCA 1746
 Db 530 PheIleAsnIleHisTrpGlyAsnLeuSerAlaThrGluGluGluValTrpAsnAlaAla 549
 QY 1747 AAATTAAGCTGACCTCATGATGCAATCTTTCGAATGCCACATGATATGACACCAAGTA 1806
 Db 550 ArgArgAlaAlaIleHisAspThrIleMetLysPheProAspLysTrpSerThrAlaVal 569
 QY 1807 GGGGAACGAGAGCTCAAGCTTTCAGAGAGGAAAGCAAGATGCAATTCAGAAAGCC 1866
 Db 570 GlyGluArgGlyLeuMetLeuSerGlyGlyGluGluGluValAlaLeuAlaAlaArgAla 589
 QY 1867 ATTTTGAAGACCCCAAGCTCATATGATGATGAAGTCACTTCACTGATGATGATGATGAT 1926
 Db 590 PheLeuLysSerProAlaIleLeuLeuCysAspGluAlaThrSerAlaLeuAspSerLys 609
 QY 1927 ACTGAAGAGCTATCTTGGTGCATGAGATGATGATGATGATGATGATGATGATGATGAT 1986
 Db 610 ThrGluAlaGluIleMetLysThrLeuArgSerLeuAlaSerAsnArgThrCysIlePhe 629
 QY 1987 ATTGCACACAGATTGTCACAGATGTTGATGACAGATGAATCATTTGCTTGATCAAGGT 2046
 Db 630 IleAlaHisArgLeuThrThrAlaMetGlnCysAspGluIleLeuValMetGluLysGly 649
 QY 2047 AAGTACCGCAGAGTGGTACCCCAATGCTTGGCTTGAACCTCATATGATCTATTTCA 2106
 Db 650 LysValValGluLysGlyThrHisGluValLeuLeuGlyLys--SerGlyArgLysAla 668
 QY 2107 GAAATGCGCATACACAGACAGCCGCTGTGACG 2139
 Db 669 LysLeuTrpThrGlnGlnAsnSerLysLeuGlu 679

RESULT 3
 T39154
 A:probable ABC transporter - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #ext_change 06-Oct-2000
 C:Accession: T39154, T37701
 R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21748
 A:Accession: T39154
 A:Molecule type: DNA
 A:Residues: 1-624 <OL1>
 A:cross-references: EMBL:Z99168, PIN:CA16305.1, GSPDB:GN00066, SPDB:SPAC89.18
 R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1997
 A:Reference number: Z21738

A:Accession: T37701
A:Molecule type: DNA

A:Residues: 227-693 <MUR>

A:Cross-references: EMBL:Z97208; PIDN:CA810098.1; GSPDB:GN00066; SPDB:SPAC15A10.01

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Alignment Scores:

Pred. No.:	1,49e-103	Length:	693
Score:	1538.50	Matches:	290
Percent Similarity:	69.50%	Conservative:	127
Best Local Similarity:	49.50%	Mismatches:	172
Query Match:	36.17%	Indels:	11
		Gaps:	4

AF133659 (1-2345) x T39154 (1-693)

```

Qy 337 AAAGATGTTGATCTCGGAAATCATAAAGCATGCTTTCTTATGTGCCCCAAAGAC 396
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 LysaenVal1ThrGlyGlnMet1LeVal1LysaPmet1LeuGlnTyr1LeTyrProLysGly 111
Qy 397 AGGCCAGATCTACGAGCTAGAGTTCATTTGCTGGGATTTTGGGTGTCAAAGGCC 456
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 112 LysThrAsnLeuLysVal1ArgVal1Ser1AlaLeu1AlaLeuVal1Ala1Lys1Ile 131
Qy 457 ATGAATATGTGCTCCCTTCATGTTTAAATGCTGTAGACAGCCCTCAACGAGTGTG 516
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 132 LeuAsnVal1GlnVal1ProPheTyrPheLysSer1Ile1LeuThrMetAsn----- 148
Qy 517 GGAACATGCTGACCTGAGTATGACCAATATACATTGCAACCATGCGAACACAGATT 576
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 149 -----ThrThrLeuVal1Gln1Lys1Val1Gly1AlaLeuTyrPseThrVal1Gly1Val1 165
Qy 577 CTGATGGCTATGCTGATTCAGAGCTGGAGCTGCTTTTAAAGATGCGAATATGCA 636
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 166 ValLeuGlyTyrGlyPhe1AlaArg1LePheSerThrVal1PheGlnGlnLeuArgAsnSer 185
Qy 637 GTATTGGCAAGTAGCCCAAGATTCATCCAGAAATAGCCCAAAATGCTTTTCTCCAT 696
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 186 ValPheAla1IleVal1SerGlnSer1Ala1LeuArgSerValSerSerAsnValTyrGlnHis 205
Qy 697 CTTCAAACTTGATCTGGGCTTTTCACTGAGACAGACAGACGGAGCTTTATCTAAGCT 756
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 206 LeuLeuAsnLeuAspMetAsnPheHisLeuSerLysGlnThrGlySer1LeuThrArg1Ala 225
Qy 757 ATTGACAGAGGAACAGGGTATCAGTTTTCCTCGAGAGCTTTGGTATTATTAATCTTCT 816
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 226 MetAspArgGlyThrLysGly1LysSerPhe1LeuSerSerMetValLeuHis1Ile1Ile 245
Qy 817 CCCATCATGTTTGAAGTATGCTTGTCACTGTGTTTGTATTATTAACAAATGCCGTCCGAG 876
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 246 Pro1LeuThrLeuGln1Leu1AlaMetVal1SerGly1LeuThrTyrLysTyrGlyProSer 265
Qy 877 TTTGCTTTGGTAACTTGGTGAACACTGTGTACATACACAGCATTCACAGTTGCAGTACA 936
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 266 PheSer1Ala1IleAla1AlaThrThrVal1AlaLeuTyrAlaLeuPheThrVal1ArgThrThr 285
Qy 937 CGGTGAGAGACATGATTTTGAATAGAAATGACAAAGACAGAAATGAGTGCAGTATAGT 996
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 286 SerTyrPArgThrVal1PheArgArgGln1AlaAsn1Ala1AspSerLys1AlaSer1Ala1Ala 305
Qy 997 GCTATAGACTCACTGCTGAATATGAAATCTGAAGTATTTTAAATGAATGAAGATATGAA 1056
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 306 Ala1IleGlnSerLeu1Leu1Ser1Gly1Lys1Val1Lys1ThrPheAsnAsnLysSer1Gly 325
Qy 1057 GCACAGATATGATGATTTTGAAGACGATATGAGACTGCTTCATTTGAAAAGTACTCT 1116
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 326 MetSerArgTyrGlnLysHis1LeuSer1AlaTyrGlnLys1AlaAsnVal1LysVal1AlaSer 345
Qy 1117 ACTGCGCTATGCTGAACCTTGCAAAATGCTATTTACAGTGTGCGTTTACAGCTATA 1176
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 346 Ser1Leu1AlaPheLeuAsnSer1Gly1Ala1Ala1Ile1LeuPseThr1AlaLeuThrLeuMet 365
Qy 1177 ATGTGCTCGCCAGTCACGGAATTTGCGAGGTACCTTACGTTTGAGATCTAGTAATG 1236

```

```

Db 366 MetTyrMetGlyCyAsArgGly1LeuVal1ThrSerAsnLeuThrVal1GlyAspLeuValMet 385
Qy 1237 GTGATGAGACAGCTTTTACACTTTTACCTTACCTCGAAGCTTTTGGGAACTGTATATGA 1296
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 386 IleAsnGlnLeuVal1PheGlnLeuSer1LeuProLeuAsnPheLeuGlySerValTyrArg 405
Qy 1297 GAGACTAGACAAAGCACTCATGATATGAAACACTTGTATTACTTACTCAAGTAGACACC 1356
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 406 GluMetArgGln1AlaPheThrAspMetGln1LeuPheSerLeuLysArg1LeuAsn1Ile 425
Qy 1357 CAATTTAAAGCAAAAGCATGGCATCTCCCTTACAGATCACACACAGACAGCTACCGTG 1416
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 426 GlnValLysGln1AlaProAspAlaArgAspLeuVal1Leu-----LysGlyLysSer1Le 443
Qy 1417 GCGTTGATATGTCATTTTGAATATACATTGAGGCGCCAGAAAGCTTTAGAGATATCC 1476
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 444 GlnPheAspAsnVal1HisPheSerTyrAsnProAsnArgPro1LeuAsnGlyCyAsSer 463
Qy 1477 TTTGAAGTCCCTGACAGAAAGAAAGTGGCCATTTGTAGAGGTAGTGGTACGAGAAAGC 1536
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 464 PheAsn1LeuPro1AlaGly1AlaLysVal1AlaPheVal1Gly1AlaSerGlyCyAsGlyLysSer 483
Qy 1537 ACAATATGAGGCTATTTATTTGCTTCTATGAGCCCTCAAAAGGTAGCATTTATCTTCT 1596
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 484 Thr1LeuArgLeuLeuPheArgPheTyrAspThrAspSerGlyLys1Leu1IleAsp 503
Qy 1597 GGTCAAAATATACAAAGATGTAGCTGAGCGTGAAGACCTTGGAGGAGGAGGTGCTACCT 1656
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 504 AsnGlnArgLeuAspGln1IleThrLeuAsnSerLeuArgLys1Ala1IleGlyVal1AlaPro 523
Qy 1657 CAGATGCTGTCTCTTCCATATATATATTTATTAACAACCTTTATATGGAACATCAGT 1716
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 524 GlnAspThrProLeuPheAsnAspThr1LeuLysThrAsn1IleGlyTyrGlyAsnProLys 543
Qy 1717 GCTTCACTGAGAAAGTGTATGACGTGGCAAAATTAAGCTGACATTGATGCAATTCTT 1776
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 544 AlaSerAsnAspGln1IleVal1Gly1AlaLysLys1AlaLys1IleHisAsp1Ile1Glu 563
Qy 1777 CGAATGCCACATGATATGACAGCCAGTAGGGAGACGAGACGCTTACAGTTCAGAGAGA 1836
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1837 GAAAGCAAAAGATGACATATGCAAGACCATTTTGAAGACCCCGCATCATCTAT 1896
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 564 GluLysGln1ArgLeu1AlaVal1SerArgLeuLeuLeuLysAsnProGln1IleLeuPhePhe 603
Qy 1897 GATGAGCTACTTCAATGCTTATGATTCGATTACTGAAGACTATTTGCTGCCATGAAG 1956
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 604 AspGln1AlaThrSer1AlaLeuAspThrAsnThrGln1Arg1AlaLeu1ArgAsn1IleAsn 623
Qy 1957 GATGTGTCGAAA-----CAGAGAACTTCTATTTTCAATTCAGACACAGATTGCAACAGTG 2010
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 624 AspLeu1IleLysGlySerHis1LysHisThrSerVal1Phe1IleAlaHisArgLeu1ArgThr1Le 643
Qy 2011 GTTGATGACAGATGAATCATTTGCTGTGATCAGAGGTAAAGTAAAGCCGATACCCAC 2070
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 644 LysAspCysAsp1Ile1LeuPheVal1LeuGlnLysGlyArgVal1GlnGlnGlySerHis 663
Qy 2071 CATGGTTGCTTGTAAACCTCATAGTATCTATTTCAGAAATGTGGCATACAGACAGC 2130
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 664 GluGlnLeuMet1AlaLys--AsnSerValTyrThrSerMetTyrHisSerGlnLysSer 682

```

RESULT 4

S54211
ATM1 protein precursor - yeast (Saccharomyces cerevisiae)
N:Alternate names: MYU protein; protein YM952.03c; protein YMR301c
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence revision 01-Sep-1995 #text_change 02-Feb-2001
R:Accession: S54211, S53971, S51801
R:Kispaal, G.; Lill, R.; Neupert, N.
submitted to the EMBL Data Library, November 1994
A:Reference number: S54211

[illegible][illegible]

Db 463 AspThrValLeuPheAenAspThrLeuArgTyrAsnIleGlyTyrGlyLeuThrAspSer 482
Qy 1720 TCACCTGAGGAAGTGTATGACGTGGCAAAATTAAGCTGATTCATGATGCAATTCCTGCA 1779
Db 483 ThrAspGluGluIleGluArgAlaIleAlaLeuAlaHisIleHisGluPheIleIleSer 502
Qy 1780 ATGCACATCGATATGACACCCAGTAGGGAGAGACATCCAGCTTCAGAGAGAA 1839
Db 503 LeuProAspGlyTyrGluThrAlaArgValGlyIuArgGlyLeuLeuSerGlyGly 522
Qy 1840 AAGCAAGAGTAGCAATTCAGAACGCCATTTTGAAGAGCCCGCCAGTCATCTCATATG 1899
Db 523 LysGlnArgValAlaIleAlaHisGThrIleLeuLysProSerIleLeuValPheAsp 542
Qy 1900 GAACCTATCTCATGCTTATGATTCGATTAATGACAGACATATCTTGTGCGCATGAGAT 1959
Db 543 GluAlaThrSerAlaLeuAspThrHisThrGluArgGluIleGlnSerHisLeuArgGlu 562
Qy 1960 GTGGTCAAAACACAACTTCATATTTTCATTCACACAGATTTGCAACGTGGTTGATGCA 2019
Db 563 ValSerArgAspHisThrThrLeuValIleAlaHisArgLeuSerThrIleIleAspAla 582
Qy 2020 GATGAATCATGTCTTGATGACAGGTAGAGCCGAGCGTGAACCCAGTCATCCATGCTTG 2079
Db 583 AspGluIleIleValIleGluAlaGlySerIleValGluArgGlyArgHisAspGluLeu 602
Qy 2080 CTTCCTAACCCCTCATAGTATCTATTCAGAAATGTGGCATACAG 2124
Db 603 LeuLeuLys---AenGlyArgTyrAlaGlyMetTrpGlnAenGln 616
RESULT 6
G71731
mitochondrial transporter atml precursor (atml) RP205 - Rickettsia prowazekii
C/Species: Rickettsia prowazekii
C/Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C/Accession: G71731
R/Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Bonten, T.; Alsmark, U
Nature 396, 133-140, 1998
A/Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A/Reference number: A71630; MUID:99039499; PMID:9823893
A/Accession: G71731
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-609 <RAND>
A/Cross-references: GB:AJ235270; GB:AJ235269; NID:G3860572; PIDN:CA14670.1; PID:e134251
A/Experimental source: strain Madrid E
C/Genetics:
A/Gene: atml; RP205
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C/Keywords: ATP; mitochondrion
F/372-566/Domain: ATP-binding cassette homology <ABC>
Alignment Scores:
Pred. No.: 6.83e-97 Length: 609
Score: 1446.50 Matches: 294
Percent Similarity: 68.67% Conservative: 118
Best Local Similarity: 49.00% Mismatches: 170
Query Match: 34.00% Indels: 18
DB: 2 Gaps: 6
AF133659 (1-2345) x G71731 (1-609)
Qy 370 ATGCTTCTATGTGTGGCCAAAGACAGACAGATCTAGAGCTAGATGTCATTTG 429
Db 23 LeuLeuThrTyrIleuHisPheSerLysAspPhe---AspIleArgLeuAlaGlyIleValIleThrSer 41
Qy 430 CTGGAGATTTTGGGTGTGCAAAAGCCATGAATATTTGGTGTCTTCCTTCATGTTAAATAT 489
Db 42 LeuIleCysLeuValIleAlaLysValIleAsnIlePheValProIleValTyrLysTyr 61
Qy 490 GCGTATAGACGCTCAACGACGATGTGCGGAAACATGCTGAACCTGAGTGAATGACCAAT 549
Db 62 IleIleAspGlyLeuAenGln-----AsnLeuSer----- 71

Qy 550 AAGTTGCAACATGCGCAACAGAGTTCTGATTTGGCTATGCTGATTCAGAGCTGAGCT 609
Db 72 -----LeuSerValLeuIleGlyValIleIleGlyTyrGlyGlyThrLysIleLeuAla 89
Qy 610 GCTTTTTCACCAAGTTCGAAATGCAAGTATTTGGCAAGTAGCCCAATTCATTCGGA 669
Db 90 GlnIlePheSerGluLeuArgAsnIleIlePheSerLysValGlyCysGlnAlaThrArg 109
Qy 670 AGAATAGCCAAAATGCTTCATCTTCATCTTCACAACTCGATGCGGGATTTTCACCTGAGC 729
Db 110 LeuValAlaLeuAsnValIlePheLysHisMetHisAsnLeuSerMetArgPheHisIleThr 129
Qy 730 AGACAGCGGAGCCTTATATGAAAGCTATTTAGACAGAGAAACAAGGGTATCAGTTTTC 789
Db 130 ArgLysThrGlyGlyLeuSerArgSerIleGluArgGlyThrLysGlyIleGluAlaVal 149
Qy 790 CTGAGTCCTTGGTATTTAATCTTTCCATCATGTTTGAAGTGAATGATGCTTTCAGTGT 849
Db 150 LeuArgTyrSerLeuPheAsnIlePheProThrSerLeuGluIleIleLeuValIleGly 169
Qy 850 GTTTGTATTACAAATGCGCGTGCAGTTTGCTTGGTAACTTGGAACTTGTGATCA 909
Db 170 IleLeuTrpTyrPheHisGlyIleTrpPheAlaValThrLeuLeuIleTrpMetIleVal 189
Qy 910 TACACAGCATTCACAGTTGACGTACACGCGTGAGAACTGATTTGAATGAATGAAC 969
Db 190 TyrValCysTyrThrLeuLeuIleSerThrTrpArgIleSerPheAlaArgGluMetAsn 209
Qy 970 AAAGCATATATGATGACAGGTAATGCTGCTATAGACTCACTGCTGAATTAATGAACCTGG 1029
Db 210 GlnSerAspAsnThrAlaAsnMetArgAlaIleAspSerLeuLeuAsnPheGluThrVal 229
Qy 1030 AAGTATTTTATATGAAGATATGAAAGCATGAGACAGAGATATGATGATTTTGAAGACGTAT 1089
Db 230 LysTyrPheAsnMetGluGluTyrGluAlaIleLysPheAsnMetAlaLeuGlnAlaTyr 249
Qy 1090 GAGACTGCTTCATGGAAGATGAACTCTACTGCTGCTATGCTGAACCTTGGTCAAGTGT 1149
Db 250 GluLysSerAlaThrLysIleThrAsnSerLeuSerIleLeuAsnIleGlyGlnAspVal 269
Qy 1150 ATTTTCAAGTGTGCTTTTAACAGCTATTAATGCTGCTGCCACAGTCAGGAATTTGGCAGGT 1209
Db 270 IleIleSerLeuGlyLeuValSerLeuMetIleLeuSerValAsnAlaIleAenGlnAsn 289
Qy 1210 ACCCTTACGTTGGAGATCTAGTAATGAGTAATGACACTGCTTTTACGTTTATTAACC 1269
Db 290 LysMetMetValGlyAspLeuIleMetValAsnAlaTyrLeuPheGlnLeuSerIlePro 309
Qy 1270 CTGAACCTTCTGGAACCTGATATTAAGAGACACTAGACAAGCACTCATATGATGAACACC 1329
Db 310 LeuSerIleLeuGlyPheAlaTyrArgGluIleLysAsnAlaLeuValAsnMetGluAsp 329
Qy 1330 TTGTTTACTCTACTCAAGGTAGACACCAATTTAAAGACAAATGATGGCATCTCCCTTT 1389
Db 330 MetPheLysLeuLeuAspIleProAlaGluIleGlnAspSerValAspAlaLysGluLeu 349
Qy 1390 CAGATCAACACAGACAGCTACCGCGCTTGTGAATATGATGATTTTGAATTCATTTGAG 1449
Db 350 IleIleSer-----LysCysLysValSerPheAsnAsnValSerPheAlaTyrAsnLys 367
Qy 1450 GGCCAGAAAGTCTTATGTAATATCTTTGAAGCTCCGACAGAAAGAAAGGCGCATT 1509
Db 368 GluArgThrIleLeuHisAsnIleTrpPheThrIleGluSerClyLysThrIleAlaVal 387
Qy 1510 GTAGAGGTAGTGGGTGACGGGAAAGACAAATAGTAGGCTATTTATTTGCTTATGAG 1569
Db 388 ValGlySerSerClyAlaIleGlyLysSerThrIleSerArgLeuLeuPheArgPheTyrAsp 407
Qy 1570 CCTCAAAAGGTAGCATTTATCTGCTGTCAAAAATATCAAGATTTGAGCCTCGGAAAGC 1629
Db 408 IleAsnSerClySerIleIleAspAsnGlnAspIleArgGluValLysGlnGlySer 427

Db 384 PheaspValProIaGlyLysThrValAlaIleValGlyProSerGlyAlaGlyLysSer 403
 1537 ACAATAGTGGCTATTATTGCTTCTATGAGCCTCAAAAGGTAGCATTTATCTTGGT 1596
 404 ThrIleSerArgLeuPheArgPheThrAspIleGlnSerGlySerIleThrIleAsp 423
 1597 GGTCAAAATATACAAAGATGTAGCCTGGAAAGCCTTGGAGGAGCGAGGTGTACTT 1656
 424 GlyIleAspValArgAspValThrGlnIleSerLeuArgLysPheIleGlyMetValPro 443
 1657 CAGATGCTGCTCTCTCAATATATCTATTATCAACCTCTTATTTAGAAACATCAT 1716
 444 GlnAspThrValIleuPheAspThrIleAlaIleAlaArgThrIleuLysAlaProProlIleuIleu 463
 1717 GCTTACCTGAGGAAGTGTATGACGAGCAAAATAGCTGACCTTCAATGCAATTTCT 1776
 464 AlSerGlnIleAspValGlnLysAlaIleGlnLeuAlaGlnIleAlaGlyPheIleLys 483
 1777 CGAATGCCACATGATATGACACCCAAAGTAGGAGACGAGACTCAAGCTTTCAGAGCA 1836
 484 HisLeuProAspGlyTyrLysSerMetValGlyIleArgGlyLeuLysLeuSerGlyGly 503
 1837 GAAAGGAAAGAGTAGAATTTGCAAGGCCATTTGAAGACCCCGAGTCATCTCTAT 1896
 504 GlnLysGlnArgValAlaIleAlaIleAlaArgThrIleuLysAlaProProlIleuIleu 523
 1897 GATGAAGCTATCTCATGCTTATGATTCATTAAGATTCATTAAGATTCATTCATTCAT 1956
 524 AspIleuAlaThrSerAlaIleuAspThrAlaThrGlnGlnIleGlnSerAlaIleuAsp 543
 1957 GATGTGTCAACACAGAACTTCTATTCTTCAATTCACACAGATTTCAACAGTGTGAT 2016
 544 IleValSerArgLysArgThrThrLeuValIleAlaHisArgLeuSerThrValIleGly 563
 2017 GCAGTGAATCATCTGCTTGGATGACGAGGTAAGGTAAGCCGAGCTGTATCCACATGCT 2076
 564 AlaAspIleuIleValIleuLysAspGlyLeuIleAlaGlnArgLysThrHisArgHis 583
 2077 TTGCTTCTAACCCTCATATGATTCATTCAGAAATGTGCATACACAGAGCCGTGTG 2136
 584 LeuLeu---AspGlnLysGlyLeuTyrAlaSerMetTrpAspArgGln----- 598
 2137 CAGAACCATGATTAACCCCAATGGAAGCAAGAAATATATATCAAAAGAGAGGAA 2196
 599 -----ArgGlnIleAspSerGlnIleAlaGlnArg 607
 2197 AGAAGAACTACACAGAA 2214
 608 LeuArgGlnValArgGln 613
 RESULT 8
 AG2707
 hypothetical protein Atcu1064 [imported] -"Agrobacterium tumefaciens (strain C58, Dupont)
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002
 C/Accession: AG2707
 R/Wood, D.W.; Secubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.
 erage, G.; Giller, W.; Grant, C.; Gunthner, D.; Kutnyan, T.; Levy, R.; Li, M.; McClell
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, B.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A/Reference number: AB2577, PMID:11743193
 A/Accession: AG2707
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-628 <RUR>
 A/Cross-references: GB:AB008688; PDB:AA142077.1; PID:gl7739457, GSPDB:GN00186
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: Atcu1064
 A/Map position: circular chromosome

C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 Alignment Scores:
 Pred. No.: 1,04e-91 Length: 628
 Score: 1375.00 Matches: 281
 Percent Similarity: 65.36% Conservative: 119
 Best Local Similarity: 45.92% Mismatches: 190
 Query Match: 32.32% Indels: 22
 DB: 2 Gaps: 4
 AF133659 (1-2345) x AG2707 (1-628)
 379 TATGTGTGCCCCAAGACAGCCGACGATCTACGAGCTTGCATTCCTGGGATTT 438
 25 TyMetTrpProGlnGlyArgTrpAspLeuLysMetArgValIleTrpAlaThrValPhe 44
 439 TTGGGTGTGCAAGGCGCATGATTTGGTCCCTCATGTTTAATATGCTGTAC 498
 45 LeuIleValAlaLysLeuValIleuIleAlaValAlaProThrPheLeuLysTrpAlaThrAsp 64
 499 AGCCTCACCCAGATGTGCGGAAACATGCTGAACCTGATGATGACACCAATATACATTGCA 558
 65 AlaLeuAsn-----GlyLysLeuAspMetAlaIleLysValProAlaPheLeu 80
 559 ACATGCGAACACAGACTTGTATGCTGATGCTATGATCAAGACCTGGAGCTGCTTTT 618
 81 LeuGlyAlaValAlaLeuValIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 100
 619 AAGGAATTCGAATTCAGATTTGGCAAGTACCCGAAATTCATCCGAATAGCC 678
 101 AsnGlnLeuArgAspSerLeuPheAlaSerValGlyGlnHisAlaValArgGlnLeuAla 120
 679 AAAAATGCTTCTTCATCTTCAACCTGATGATGCTGGGTTTCACTGACGACAGACAGC 728
 121 TyrArgThrPheValHisMetHisArgLeuSerLeuArgPheHisLeuGlnArgLysThr 140
 739 GAGCTTATCTAAGCTATTTGACAGAGAAAGGGTATCATGTTTGTCTGAGTGT 798
 141 GlyLysLeuSerArgValIleGlnArgGlyThrLysGlyIleGlnThrIleValArgPhe 160
 799 TTGGTATTAATCTTCTCCATCATGTTGAAGTATGCTGTGAGTGTGTGTAT 858
 161 ThrIleLeuAsnThrAlaProThrPheIleGlnPheLeuLeuThrAlaIleIlePheTrp 180
 859 TACAATGCGCGTCCAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 918
 181 AlaSerTyrGlyPheSerTyrValIleuValThrValIleValIleValIleValIleVal 200
 919 TTGACAGTTGACGCTACACGCTGAGAGAACTAGATTAGATAGAAATGAAAGAGAT 978
 201 PheThrValArgAlaSerAsnTrpArgIleGlyIleArgArgAlaMetAsnAspSerAsp 220
 979 AATGATGACGATATGCTGTATAGTCACTACGCTGATATGAAATGAACTGGAAGTATTT 1038
 221 ThrAspAlaAsnThrLysAlaIleAspSerLeuLeuAsnPheGlnThrValLysTyrPhe 240
 1039 AATTAAGAAAGATATGAGACAGAGATAGATGATTTTGGTGAAGCGTATGAGAGTGT 1098
 241 GlyAsnGlnGlnMetGlnAlaArgArgPheAspValAlaMetGlnArgLysLysSer 260
 1099 TCATTGAAAAAGTACTCTATCTGCTGCTATGCTGAATTTGGTCAAAAGTGTATTTCACT 1158
 261 AlaIleSerIleTrpThrSerLeuGlyTrpLeuAsnPheGlyGlnGlyValIlePheGly 280
 1159 GTTGGTTTAAACACTTAATGTTGCTGCCAGTACAGGAATGTGACAGTACCTTACT 1218
 281 IleGlySerThrIleMetMetValMetSerAlaLeuAlaValGlnArgLysGlnGlnThr 300
 1219 GTTGGATCTATGATATGAGATGAGAGCTTTTTCAGCTTTCATACCCCGAAGCTTT 1278
 301 IleGlyAspPheValPheValAsnAlaLeuLeuLeuGlnIleuSerValProLeuAsnPhe 320
 1279 CTGGGAACCTGTATATAGAGAGACTAGACAAAGCATCATGATATGAAACCTTGTACT 1338

[illegible]

RESULT 9
H97489
mitochondrial transporter atmi precursor (atmi) rp205 [imported] - Agrobacterium tumefaciens
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: H97489
R/Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A:Reference number: A97359; PMID:11743194
A:Accession: H97489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-629 <KOR>
A:Citations-References: GB:AE007869; PIDN:AAK66873.1; PID:915156091; GSPFD:GN00165
C:Genetics:
A:Gene: AGR_C_1966
A:Map position: circular chromosome

[illegible]

QY	379	TATGTGTGGCCAAAGACAGGACGAGTCTTCAACAGAGCTAGAGTTGGCCATTTCGCTGGGATTT	438
Db	26	TYRMECTPRGUGUGLYATGTPRASPRLUYSMECAVALIETRLPRLATHVALPHE	45
QY	439	TTGGGTGGTGCAAAGGCGCATGAATATTGTGGTTCCTTCATGTTTAAATATGCTGTAGAC	498
Db	46	LEUILEVALALALYBLEUVALLEUILEALVALPROTYRPHENELYSTRLPALHTRSP	65
QY	499	AGCCTCAACGACGATGTGGGAAAATGCTGAACCTGAGTGATGACCAAAATACAGTTGCA	558
Db	66	ALALEUAMEN-----GLYBLEUASPMETALGLYLEUVALPROALAPHELEU	81
QY	559	ACCATGGCAACAGCATTTCTGATTCGCTATGCTGATCTCAAGAGCTGGACGCTCTTTTTT	618
Db	82	LEUGLYLALVALALALEUVALIIEALATYRSMLEUHTHRARGLEUILEGLNALGLYLEU	101
QY	619	AACCAAGTTGCAAAATGACGATATTGGCAGGTRAGCCCAAAATTCATCCGAAGATATACC	678
Db	102	ASNGINLEUARGAPSERLEUPHEALASERVALGLYGLNHLALVALARGLEULEUOLA	121
QY	679	AAAAATGCTTTTCCATCTTCCATCCCAACCTGGATCTGGGTTTCCACTGACGACAGACG	738
Db	122	TYRARGTHRPHEVALHLSWECHSHARGLEUSERLEUARGPHEHISLEUCLINARGYSTHR	141
QY	739	GGAGCTTATCTAAGCGCTATTGACAGAGCAACAGGGGATACAGTTTTCCTCGATGCT	798
Db	142	GLYGLYLEUSERARGVALIEGLUARGLYTHIRLYGLYILEGLUTHIRLEVALARGPHE	161
QY	799	TTTGATTTTATCTTCTTCCCATCATGTTGAAGTGATCTTGTCAGTGGTGTTTGAT	858
Db	162	THRILEUAMENHTHRLAPROTHRPHEILEGLUPHEULEUHTHRLALIELEPHEPTR	181
QY	859	TACAAATCCGCGTCCCGCATTTGCTTGTTGTRACCTTGGAACCTTGTRACATACACACA	918
Db	182	ALASERTYRGLYPHESEPTYRVALLEUVALTHVALTHVALTHVALTRPALARYLLETTP	201
QY	919	TTTCACAGTTGAGTCACACGCGTGGAGAACCTAGATTAGATGAATGAACAAACACAGAT	978
Db	202	PHETHRIVALARGALASERANTPRAGILEGLYLEARGARGALAMECAHSPERASP	221
QY	979	AATGATGACGGTATGCTGCTATAGACTACTGCTGAATTTGAATGAACTGTGAAGATTTT	1038
Db	222	THRASPALASANTHRLYSALALIEASPERLEULEAMSPHEGLUTHRYVALYSTYRHE	241
QY	1039	AATATGCAAGATGTAGACAGACAGATATGATGATTTTGAAGACGTATGACGACTCT	1098
Db	242	GLYASNGILUGLUMECTGLUALAARGARGPHEASPVALLAMEGLIARGYRGLIULYSER	261
QY	1099	TCATTGAAAATACCTACTACTGCTAGCTAGCTGAACCTTTGGTCAAGTGCATATTTACGT	1158
Db	262	ALALIESERIETPRMHSERLEUGLYTTPRLEAMSPHEGLYGLNGLYVALIIEPHEGLY	281
QY	1159	GTCGGTTTAAACGCTATATATGCTGTGCCAGTCAGGGAATTGTGGCAGATACCTTACT	1218


```

Oy 634 GCAGATTGGCAGATGACCCAGAAATTCATCCGAAAGATAGCCAAAATGCTTTC 693
      ::::: ||| ||| :::: ||| :::: |||
Db 332 PheLeuTrpIleArgValGlnGlnPheThrSerArgGlyValGluLeuArgLeuPheSer 351
Oy 694 CATCTTCAACAACCTGGATCTGGGTTTTCACCTGACAGACAGACGGAGCTTATCTAAG 753
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 352 HisLeuHisGluLeuSerLeuArgTrpHisLeuGlyArgArgThrGlyGluValLeuArg 371
Oy 754 GCTATTGACAGAGAACAGGGGATCAGTTTGTCTGAGGCTTGGTATTATTACTT 813
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 372 IleValAspArgGlyThrSerSerValThrGlyLeuLeuSerThrLeuValPheAsnIle 391
Oy 814 CTCCCATCATGTTTGAAGTATGATGCTGACGTGGTGGTATTATTAACAATGCG----- 867
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 392 IleProThrLeuAlaAspIleIleIle-----GlyIleIleIleThrPheSerMetPhePhe 409
Oy 868 GGTGCCAGTTTGTCTTGGTAAACCTTGACACCTTGGTACATACACAGCATTCACAGTT 927
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 410 AsnIleATrpPheGlyLeuIleValPheLeuGlySerMetSerLeuThrLeuIleLeuThrIle 429
Oy 928 GCAGTCACACGGTGGAACATGATTGAAATGAAATGAAACGACATATATATGCA 987
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 430 MetValThrGluTrpArgAlaIleAspPheArgAspMetAsnThrGlnGluAsnIleThr 449
Oy 988 GGTATGCTGCTATGATGACTGCTGATGATGAACTGAGATATTTTAATATGAA 1047
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 450 ArgAlaIleArgIleValAspSerLeuLeuAsnPheGlnThrValIlePheIleAsnIleGlu 469
Oy 1048 AGATATGAGACACAGATATGATGATGATTTTGAAGACGTATGACATGCTTTCATGAAA 1107
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 470 GlyTrpGluLeuGlnArgValArgGluAlaIleLeuGlyPheGlnGlyLeuGluTrpIle 489
Oy 1108 AGTACTCTACTCTGCTATGCTGTAACCTTGGTCAAGGCTATTTTTCAGTCCGTTA 1167
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 490 SerThrIleAspSerLeuValLeuLeuAsnGlnThrGlnAsnMetValIleGlyPheGlyLeu 509
Oy 1168 ACAGCTTATATGATGCTGCCCGCAGTACAGGATTTGGCAGTACCTTACTCTTGAAGT 1227
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 510 LeuAlaGlySerLeuLeuGlyAlaIleThrPheValSerGluArgLeuGlnValGlyAsp 529
Oy 1228 CTAGTAATGTAATGATGACTGCTTTTTCAGCTTTCATTAACCCCTGAACTTTCGGAAT 1287
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 530 PheValLeuPheGlyThrThrIleThrGlnLeuIleuThrPheLeuAsnTrpPheGlyThr 549
Oy 1288 GTATATAGAGACTAGACAGACACTCATAGTATGAAACACTTGTACTTACTTCAAG 1347
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 550 TyrTrpArgMetIleGlnThrAsnPheIleAspMetGluAsnMetPheAspLeuLeuLys 569
Oy 1348 GTAGACACCCCAATTAAGACAAAGATGATGCGATCTCCCTTCAGATCACACACAGACA 1407
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 570 GluGluThrGluValIleAspValProGlyAlaGlyProLeuArgPhe-----HisLys 587
Oy 1408 GCTACCTGCGCTTGTATATGATGCTTGTGATATCATTTGAGGGCAGAAAGTCTTAACT 1467
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 588 GlyArgValGluPheGluAsnValHisPheSerThrAlaAspGlyArgGluThrLeuGln 607
Oy 1468 GGAATATCTTTGAAGTCCCTCGACAGAAAGAAAGTGGCCATTGTGAGAGTGTGGGTCA 1527
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 608 AspValSerPheThrValMetProGlyGlnThrValAlaIleValGlyProSerGlyAla 627
Oy 1528 GGGAAAGACAGATGATGAGCTATTTATTTGCTTCTATGAGCCCAAAAGGTGACATT 1587
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 628 GlyLysSerThrIleLeuArgLeuLeuPheArgPheThrAspIleSerSerGlyCysIle 647
Oy 1588 TATCTTGTGCTCAAAATATACAAAGTGTGACCTGGAAGGCTTCGAGGGCAGATGGGA 1647
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 648 ArgIleAspGlyGlnAspIleSerGlnValThrGlnIleSerLeuArgSerHisIleGly 667
Oy 1648 GTGTAAGCTCAGAGATGCTGCTCTTCCATTAATACTATTATTATTAACAACCTTTAATGGA 1707
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 668 ValValProGlnAspThrValIlePheAsnAspThrIleAlaAsnAsnIleArgTrpGly 687

```

```

Oy 1708 AACATCAGTCTTCACCTCGAAGAGTATGACAGTGGCAAAATTAAGCTTCATGAT 1767
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 688 ArgValThrIleArgIleAspSerGluIleGlnIleAlaIleGlnIleAlaGlyIleHisAsp 707
Oy 1768 GCATTTCTTGAATGCCATGATGATATGACACCCCAAGTAGGGGAACGAGACATTAACCTT 1827
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 708 AlaIleLeuSerPheProGluGlyThrGlnIleValGlyGluArgGlyLeuLeuValLeu 727
Oy 1828 TCAGAGAGGAAAGAACAAAGAGTAGCAATTTGACAGACCAATTTGAAGACCCCGACGT 1887
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 728 SerGlyGlyGlyLysGlnArgValAlaIleAlaArgThrIleLeuValAlaProAspIle 747
Oy 1888 ATACTATGATGAGAGCTACTTCTCATGCTTATGATTCGATTCGAGAGACATTTCTTGT 1947
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 748 IleLeuLeuAspGlnIleAlaIleThrSerAlaLeuAspThrSerAsnGluArgAlaIleGlnAla 767
Oy 1948 GCCATGAAGATGTGGTCAACACAGAACTTCTTATTTTCATTGACACAGATTGTCAACA 2007
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 768 SerLeuAlaLysValCysThrAsnArgThrThrIleValAlaIleHisArgLeuSerThr 787
Oy 2008 GTGGTGAATGACAGATGAAATCATGTTGTGGATCAGGGTAAAGTACGAGACGTAAC 2067
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 788 ValValAsnAlaAspGlnIleLeuValIleLysAspGlyCysIleIleGluArgGlyArg 807
Oy 2068 CACCATGGTTTGTCTGCTTAACCTCATGATGATCTTATTCAGAAATGTGGCATACAG 2124
      ::::: ||| ||| ::::: ||| ||| ::::: ||| |||
Db 808 HisGluIleLeuLeuSerArg---GlyGlyValAlaThrIleGluMetTrpGlnLeuGln 825

```

RESULT 11

```

C87412
ABC transporter, HlyB/MedA family C01314 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: C87412
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo-
n, J.; Brimble, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: C87412
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-643 <STO>
A/Cross-references: GB:AE005673; NID:G1342655; PIDN:AAK23295.1; GSPDB:GN00148
C/Genetics:
A/Gene: C01314

```

Alignment Scores:

```

Pred. No.: 1,43e-67 Length: 643
Score: 1042.00 Matches: 253
Percent Similarity: 52.07% Conservative: 99
Best Local Similarity: 37.43% Mismatches: 266
Query Match: 24.49% Indels: 58
DB: 2 Gaps: 7

```

AF133659 (1-2345) x C87412 (1-643)

```

Oy 145 TTGGGAACCGCTCGAGCTTACAGATTCAGATCTTAAAAAGTATACATGCGACAGA 204
      ||||| ||||| ||||| |||||
Db 5 LeuGlyAspAlaIleArgTrpArgThrAspHisLeuAlaAspProLysGlyTrpAlaMet 24
Oy 205 TTGGGAAGAGGCAT-----TCAGGACAGTCTTGAT 237
      ||| ||| ||| ||| ||| ||| ||| |||
Db 25 ArgGlyLeuGlyGlyProProSerProGlyArgGluAlaValAlaGlyAlaValAlaLys 44
Oy 238 GCTGCAAGGCTCTCCAGATGCGCACTGATAGAAAAAGACATGTGGCATGTCTAT 297
      ||| ::::: ||||| |||
Db 45 AlaAspGlnProIleAspPheTrp----- 52
Oy 298 GCAGAGAGAGATCTCCACACAGACCCAAAGAGGTTTAAAGATGTTGATATCTCGAAA 357
      ||||| ||||| |||
Db 53 -----Lys 53

```

QY 358 ATCAATAAGCATGCTTTCTTATGTGTGSCCAAGACAGGCCAGATCTAGAGCTAGA 417
 Db 54 AlametserAepLeuLeuGlnLeuValLeuArgSerGlnAlaProGlyLeuArgTyrArg 73
 QY 418 GTTCCATTTGGCTGGGATTTTGGGTGGTGCAGAAAGCCATGAATATTGGTTCCTTC 477
 Db 74 LeuThrValAlaLeuLeuLeuThrLeuThrGlySerValLeuGlyValLeuAlaProLeu 93
 QY 478 ATGTTTAAATATGCTGTAGACAGCTCAACAGATGTGGGAAACATGCTGACCTGAGT 537
 Db 94 MetLeuGlyGlnAlaVal-----AenGlnLeuSerValGly 105
 QY 538 GATGCACCAATATACAGTTCACCATGGCAACAGACTTCTG---ATTGGCTATGTGTA 594
 Db 106 GlnGlyValAlaValAlaThrValThrLeuAlaPheAlaSerLeuAlaIleGlyTyrAlaLeu 125
 QY 595 TCAGAGACTGGAGCTGCTTTTAAAGAAAGTTCGAATGACGATTTGGCAAGTGGCC 654
 Db 126 ValArgPheIleSerAlaAlaAlaProGlnAlaArgAspThrIlePheThrProValAla 145
 QY 655 CAGATTCGAATCCAGAGATAGCAAAATATGCTTCATCTTCACCACTGATGCTG 714
 Db 146 GlnAlaAlaGlnThrArgAlaAlaValGlnThrPheAlaHisAlaLeuSerLeuSerIle 165
 QY 715 GGTTCACCTGACGACAGACAGCGAGCTTATTAAGGCTATTGACAGAGAAAG 774
 Db 166 AspPheHisGlnSerIleValArgThrGlySerLeuSerArgValIleAspArgGlyAlaArg 185
 QY 775 GGATCATGTTTGGCTGAGTGGCTTTGATTTAACTTCTTCCCATCATGTTGAAGTG 834
 Db 186 SerMetAspPheLeuLeuArgGlyLeuValPheAsnLeuAlaProThrGlyIleGlnLeu 205
 QY 835 ATGCTTGTACGTGGTGTGTTTGTATTACAAATGCGGTGCCGATTTGGTGTACCTT 894
 Db 206 IleLeuAlaAlaValAlaLeuAlaAlaValAlaTyrAspThrPheAlaAlaValAlaLeu 225
 QY 895 GGAACACTGGTATACACACAGCATTCAGTTGACAGTCAGTCAGGAGAGAACTGATTT 954
 Db 226 ValThrValAlaIleTyrGlyTyrValThrPheAlaIleSerAspTyrArgIleGlyHis 245
 QY 955 AGAATAGAAATGAACAAGCAGATTAATGATGACGATTAAGTCTTAATGACTCATCTG 1014
 Db 246 ArgArgAlaLeuAsnGlnAlaAspAlaIleGlnAlaIleGlyArgAlaValAlaLeuLeu 265
 QY 1015 AATATGAACCTGTGAGATTTTAAATGAAGAATATGAGACAGAGATTAATGATGGA 1074
 Db 266 AsnTyrGlnThrValLysSerPheGlyGlyGlnAlaArgAlaValAlaSerTyrGlnArg 285
 QY 1075 TTTTGAAGCGTATGAGACTGCTTCAATGAAGTACTCTGCTGCTGATGCTGAAC 1134
 Db 286 AlaLeuAspThrTyrGlyArgAlaAsnIleLeuAlaThrGlnSerLeuAsnLeuLeuAsn 305
 QY 1135 TTGCGTCAAGTGTCTATTTCAGTGTGGTTTAAACACTTAATGTCGTCCAGCTGAG 1194
 Db 306 LeuValGlnSerGlyValMetSerValGlyLeuGlyValMetAlaValLeuAlaGlySer 325
 QY 1195 GGAATGTGACGAGTACCTTACTGTTGAGATCTACTAATGATGATGAGACTGCTTT 1254
 Db 326 GlnAlaAlaHisGlyArgMetGlyProGlyAspAlaThrAlaAlaValLeuIleLeuIle 345
 QY 1255 CAGCTTTCATTAACCTGGAATCTTCTGGGAACTGTATATAGAGATAGACAGCAAGCTC 1314
 Db 346 AsnLeuTyrAlaProLeuAsnIleLeuGlyPheAlaTyrArgGlnIleArgGlnSerLeu 365
 QY 1315 ATGATATGAACACTTGTATTACTCTAGAGTATGACACCAATTAAGCAAAAGTG 1374
 Db 366 IleAspMetGlnAlaMetLeuAspLeuArgArg-----GlnAlaAlaAspValAla 382
 QY 1375 ATGGCATCTCCCTTCAGATCACACACAGAGCT-----ACCGTGGCC 1419
 Db 383 AspAlaProAspAlaGlnAspLeuProCysAlaAspGlnArgGlyGlyAlaValAla 402

QY 1420 TTGATATATGTCATTTTGAATATCATTTAGAGGCCAGAAAGTCTTATGGAATATCTTT 1479
 Db 403 PheGlnAlaValSerPheArgHisGlyAlaArgSerGlnIleLeuSerGlnValSerLeu 422
 QY 1480 GAGTCCCTGCAGAAAGAAAGTGGCCATTTGTAGAGTGTGTGGTTCAGGAAAGACA 1539
 Db 423 ThrValAlaProGlySerThrValAlaIleValGlyProSerGlyAlaGlySerThr 442
 QY 1540 ATATGAGGCTATTTATTTGCTTCTATAGAGCTCAAAAGGGTATGACTTTTCTTGCTGT 1599
 Db 443 LeuValArgLeuAlaLeuArgMetIleAspProGlnSerGlyArgValThrLeuAspGly 462
 QY 1600 CAAATATACAGATGTAGCTGAGCAAGCTTCGAGAGGCACTGGAGTGTACTGAC 1659
 Db 463 TyrAspLeuLysAlaLeuLysGlnSerLeuArgArgAlaValAlaLeuValProGln 482
 QY 1660 GATGCTGTCTCTTCCATATATCTATTTATTAACACTCTTAATATGAAACATCAGTCT 1719
 Db 483 AspValAlaLeuPheAsnAspThrLeuAlaAlaAsnIleAlaPheAlaArgProGlnVal 502
 QY 1720 TCACCTGAGGAGTGTATGAGTGGCAAAATATGACTGATTCATGATGCAATCTTGA 1779
 Db 503 GlyGlnThrGlnValThrAlaAlaAlaGlnAlaIleGlnLeuGlyAspPheIleArgSer 522
 QY 1780 ATGCCACATGATATATACACCCAGTAGGGAAAGAGACTCAAGCTTTGAGAGAGAA 1839
 Db 523 LeuProGlnIleMetAlaThrLysValGlyGlnArgGlyLeuLysLeuSerGlyGlyGln 542
 QY 1840 AAGCAAGATATGACATTTGAAAGCCATTTTGAAGAGCCCCAGTCAATCTATGAT 1899
 Db 543 ArgGlnArgValGlyIleAlaArgAlaLeuLeuAlaGlnProArgValLeuIleLeuAsp 562
 QY 1900 GAAGTACTTCATCTGATGATGATTTACTGAGAGACTTTTCTTGCTGCTGATGAGAT 1959
 Db 563 GlnAlaThrSerAlaLeuAspSerArgThrGlnAlaAlaIleGlnAlaThrLeuArgLys 582
 QY 1960 GTGTCAAACACAGAACTTATTTTCAATGACACACAGATTTGTCACAGTGTGATGCA 2019
 Db 583 AlaArgAlaGlyArgThrThrLeuValAlaAlaHisArgLeuSerThrIleAlaAspAla 602
 QY 2020 GATGAATATCTTGTCTTGTATCAGGGTATGAGTGGCCGAAAGTGTACCCATGCTTGT 2079
 Db 603 AspGlnIleValValLeuArgArgGlyLysIleValGlnArgGlyArgHisGlnAlaLeu 622
 QY 2080 CTTGCTAACCTCATATGATCTATTCAGAAATGTGCAATACAGAGC 2127
 Db 623 LeuAla---AlaGlyGlyGlnTyrAlaAlaLeuTyrArgArgGlnThr 637
 RESULT 12
 S25198
 vacuolar membrane protein HMT1 - fission yeast (Schizosaccharomyces pombe)
 N/Alternate names: heavy metal tolerance protein (HMT)
 C/Speices: Schizosaccharomyces pombe
 C/Date: 28-May-1993 #sequence revision 28-May-1993 #text_change 02-Feb-2001
 C/Accession: S25198; T41583; T41582
 R/Ortz, D.F.; Kreppel, L.; Speiser, D.M.; Scheel, G.; McDonald, G.; Ow, D.W.
 EMO J. 11, 3491-3499, 1992
 A/Title: Heavy metal tolerance in the fission yeast requires an ATP-binding cassette-cy
 A/Reference number: S25198; MUID:93010938; PMID:1396551
 A/Accession: S25198
 A/Molecule type: mRNA
 A/Residues: 1-830 <ORF>
 A/Cross-References: EMBL:Z14055; NID:94971; PIDN:CAA78419.1; PID:94972
 R/Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, September 1998
 A/Reference number: Z21968
 A/Accession: T41583
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-78, 'R', 80-404 <MUR1>
 A/Cross-References: EMBL:AL011543; NID:94239667; PIDN:CAA20838.1; PID:94239668; GSPDB:G
 A/Experimental source: strain 972h-, cosmid C74
 R/Murphy, L.; Harris, D.; Wood, V.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.

QY	1984	TTATTTGACACAGATTTCTCAACAGTGGTATGCAGATGAAATCATGTTCTGGATCAG	2043
Db	878	lIeIIeIAhSlvSlvSlvCysThrIleIleThrAlaGluIleuIleIleIleuIleAsnLys	897
QY	2044	GGTAAAGTAGCCGAAACGTGTATCCACCCATGGTTGCTTGCTAAACCTCATAGTATCAT	2103
Db	898	GlyIvSIleIleGluIvRGlyThrIleIleuAspIleu---LysCysAsnGlyIuIv	916
QY	2104	TCGAAATGTGGCATACACAGACGACCGCTGTGCAGAACCATATTAACCCCAATGGAA	2163
Db	917	ThrGluMetTTPAsMetGlnSerIysSerAsnGluPvProHsIvThrIvAsnSerSer	936
QY	2164	GCAAAAGAAAGAAAATATATCCAAAGAGGAGGAGAAAGAAAGAAATACACAGAAATATGTC	2223
Db	937	IleAspLysAspAspValAsnLys-----AsnAsnAsnLysAsnAsnAspValIleIleu	954
QY	2224	AATAGTGTGAAA	2235
Db	955	AsnThrCysLys	958

RESULT 14

hypothetical protein W09D6.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: J26313
R:Gardner, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20195
A:Accession: J26313
A:Status: preliminary; translated from GE/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-772 <WIL>
A:Cross-references: EMBL:Z82078; PIDB:CA04947.1; GSPDB:GN00021; CESP:W09D6.6
A:Experimental source: clone W09D6
C:Genetics:
A:Gene: CESP:W09D6.6
A:Map position: 3
A:Introns: 24/3, 59/1, 111/3, 267/3, 320/1, 412/2, 483/2, 552/3, 738/2

Alignment Scores:

Pred. No.:	2.34e-58	Length:	772
Score:	915.00	Matches:	219
Percent Similarity:	50.72%	Conservative:	97
Best Local Similarity:	35.15%	Mismatches:	175
Query Match:	21.51%	Indels:	132
DB:	2	Gaps:	11

AF133659 (1-2345) x T26313 (1-772)

QY	349	ACICGGAAATCATATAAAAGCAATGCTTCTTATATGTCGCCCCAAGACAGCCAGATCTA	408
		: : : : : :	
Db	273	SerLysLysValVal---ThrValAlaProTyrIleTPrProThrLysSerLysSerLys	291
QY	409	CGAGCTAGAGTGGCATTTTCGCTGGGATTTTGGGTGTGCAAAAGGCATCAATATTTGTG	468
		: : : : : :	
Db	232	GlnLeuArgValValPheCysLeuPheLeuLeuIleIleGlyArgLeuIleAsnValSer	311
QY	469	GTTCCCTTCATGTTAAATATGCTGTAGACAGCCTCAACAGATGTGGGAAACATGCTG	528
		: : : : : :	
Db	312	LeuProIleLeuSerLysTyrIleValAspGluLeu-----	323
QY	559	AACCTGAGTATGACCAACAAATCA-----GTTGCACCACTGCAACGACGATT	576
		: : : : : : : : :	
Db	324	-----AlaThrProAspThrPheGlnTyrSerLysLeuLeuPheLeuAlaThrPheLeu	340
QY	577	-----CTGATTGGCTATAGTGTGTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTGCA	630
Db	341	LysPheLeuGlnGlyAsnGly-----AlaMetGlyGlyPheLeuAsnThrValArg	357
QY	631	AATGACAGTATTGGCAAGGTACCCAGAAATTCACCAAGAAATATGCAAAAATGCTCTT	690
		: : : : : : : : :	
Db	358	ThrTyrLeuSerIleLeuIleGlnIleIntyrThrThrArgGluLeuGluValGluLeuPhe	377

Oy	691	CTTATTTGACAACTCGATCTGGGTTCACCTGAGACAGACGGAGCTTTACT	750
Db	378	LYSHISLEUHSIERLEUSERLEUAGTRPHLSLEUSERAUYRHYETOLYLINALEU	397
Oy	751	AAGCCTATTGACAGAGAACACAGGGGTATCACTTTTGTCCTGAGTGCCTTGGATTAT	810
Db	398	ArgValMetAspArgGlyThr	404
Oy	811	CTTCTTCCCATCATGTTTGAAAGTAGTCTTGTCAGTGGTGTGTTGTATTACAAATGCCGT	870
Db	404	----	404
Oy	871	GCCAGTTGCTCTTGGTAACCCTTGGACACTTGTATACACAGATTCAAGTTGCA	930
Db	404	----	404
Oy	931	GTCACACGGTGGAGAATACTAGATTAGAAATGAACAAAGCACAGATTATGATGCAGST	990
Db	404	----	404

QY

Db 405 ---SerSerValAsnAsnIleLeuAsnTyrGlnValThrGluTrp----- 418

QY 1051 TATGAGCACAGAGATATGATGGATTTTGAAGACGTATGAGACTGCTTCATTGAAAGT 1110

D5 419 -----Lysine 420

0Y III ACCICACIIGCTAIGCIGAACIIIGICAAAGIGCIHIIICAGIGICGGIIPACH 11/

DD 421 GIMMABEIDENHATARNENENWBNLYBNECOTLWBNHATATEVEEYATEEOLYNELTIC TO

ΣY	11/1	0C	1A	1B	1C	1D	1E	1F	1G	1H	1I	1J	1K	1L	1M	1N	1O	1P	1Q	1R	1S	1T	1U	1V	1W	1X	1Y	1Z	1AA	1AB	1AC	1AD	1AE	1AF	1AG	1AH	1AI	1AJ	1AK	1AL	1AM	1AN	1AO	1AP	1AQ	1AR	1AS	1AT	1AU	1AV	1AW	1AX	1AY	1AZ	1BA	1BB	1BC	1BD	1BE	1BF	1BG	1BH	1BI	1BJ	1BK	1BL	1BM	1BN	1BO	1BP	1BQ	1BR	1BS	1BT	1BU	1BV	1BW	1BX	1BY	1BZ	1CA	1CB	1CC	1CD	1CE	1CF	1CG	1CH	1CI	1CJ	1CK	1CL	1CM	1CN	1CO	1CP	1CQ	1CR	1CS	1CT	1CU	1CV	1CW	1CX	1CY	1CZ	1DA	1DB	1DC	1DD	1DE	1DF	1DG	1DH	1DI	1DJ	1DK	1DL	1DM	1DN	1DO	1DP	1DQ	1DR	1DS	1DT	1DU	1DV	1DW	1DX	1DY	1DZ	1EA	1EB	1EC	1ED	1EE	1EF	1EG	1EH	1EI	1EJ	1EK	1EL	1EM	1EN	1EO	1EP	1EQ	1ER	1ES	1ET	1EU	1EV	1EW	1EX	1EY	1EZ	1FA	1FB	1FC	1FD	1FE	1FF	1FG	1FH	1FI	1FJ	1FK	1FL	1FM	1FN	1FO	1FP	1FQ	1FR	1FS	1FT	1FU	1FV	1FW	1FX	1FY	1FZ	1GA	1GB	1GC	1GD	1GE	1GF	1GG	1GH	1GI	1GJ	1GK	1GL	1GM	1GN	1GO	1GP	1GQ	1GR	1GS	1GT	1GU	1GV	1GW	1GX	1GY	1GZ	1HA	1HB	1HC	1HD	1HE	1HF	1HG	1HH	1HI	1HJ	1HK	1HL	1HM	1HN	1HO	1HP	1HQ	1HR	1HS	1HT	1HU	1HV	1HW	1HX	1HY	1HZ	1IA	1IB	1IC	1ID	1IE	1IF	1IG	1IH	1II	1IJ	1IK	1IL	1IM	1IN	1IO	1IP	1IQ	1IR	1IS	1IT	1IU	1IV	1IW	1IX	1IY	1IZ	1JA	1JB	1JC	1JD	1JE	1JF	1JG	1JH	1JI	1JJ	1JK	1JL	1JM	1JN	1JO	1JP	1JQ	1JR	1JS	1JT	1JU	1JV	1JW	1JX	1JY	1JZ	1KA	1KB	1KC	1KD	1KE	1KF	1KG	1KH	1KI	1KJ	1KK	1KL	1KM	1KN	1KO	1KP	1KQ	1KR	1KS	1KT	1KU	1KV	1KW	1KX	1KY	1KZ	1LA	1LB	1LC	1LD	1LE	1LF	1LG	1LH	1LI	1LJ	1LK	1LL	1LM	1LN	1LO	1LP	1LQ	1LR	1LS	1LT	1LU	1LV	1LW	1LX	1LY	1LZ	1MA	1MB	1MC	1MD	1ME	1MF	1MG	1MH	1MI	1MJ	1MK	1ML	1MM	1MN	1MO	1MP	1MQ	1MR	1MS	1MT	1MU	1MV	1MW	1MX	1MY	1MZ	1NA	1NB	1NC	1ND	1NE	1NF	1NG	1NH	1NI	1NJ	1NK	1NL	1NM	1NN	1NO	1NP	1NQ	1NR	1NS	1NT	1NU	1NV	1NW	1NX	1NY	1NZ	1OA	1OB	1OC	1OD	1OE	1OF	1OG	1OH	1OI	1OJ	1OK	1OL	1OM	1ON	1OO	1OP	1OQ	1OR	1OS	1OT	1OU	1OV	1OW	1OX	1OY	1OZ	1PA	1PB	1PC	1PD	1PE	1PF	1PG	1PH	1PI	1PJ	1PK	1PL	1PM	1PN	1PO	1PP	1PQ	1PR	1PS	1PT	1PU	1PV	1PW	1PX	1PY	1PZ	1QA	1QB	1QC	1QD	1QE	1QF	1QG	1QH	1QI	1QJ	1QK	1QL	1QM	1QN	1QO	1QP	1QQ	1QR	1QS	1QT	1QU	1QV	1QW	1QX	1QY	1QZ	1RA	1RB	1RC	1RD	1RE	1RF	1RG	1RH	1RI	1RJ	1RK	1RL	1RM	1RN	1RO	1RP	1RQ	1RR	1RS	1RT	1RU	1RV	1RW	1RX	1RY	1RZ	1SA	1SB	1SC	1SD	1SE	1SF
----	------	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

22 11 2017 11:11:11

.....

[illegible]

.....

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

```

nb      ||:::| ||:::| ||:::| ||:::|
519  C|v|t|t|e|s|e|r|v|a|l|y|e|s|e|r|v|a|l|e|n|t|a|u|t|h|e|r|p|h|e|g|u|t|v|r|a|s|n|t|b|r|c|j|v|e|u|p|r|o|v|a|l|t|e|l|y|e|s 538

```

1468 GGAATATTCCTTGAAGTCCCTGCAGGAAGAAAGTGGCCATTGTAGGAGGTAGTGGTCA 1527

Db 539 AsnTleSerPheGluTleGlyAsnGlyClnThrValAlaLeuValGlySerSerGlySer 558

QY 1528 GGGAAAGCACAATAGTGAGGCTATTATTTCGCTTCTATGAGCCTCAAAAGGTCAGCAT 1587

Db 559 GlySerThrLeuIleArgLeuLeuPheArgLeuPheGluSerThrGluGlySerIle 578

QY 1588 TATCTTGGTCAAAATATACAAGATGTGAGCCTGGAAGCCTTCGGAGGGCAGTGGGA 1647

Db 579 GlupheaspGlyIleaspValArgasNTyrThrmetHisSerLeuArgGInGInIleGly 598

QY 1648 GTGGTACCTCAGGATGCTGTCTCTTCATATACTATTATTACAACTCTATATGGA 1707

Db 599 ILevalProGlnsptHrvalLeupheAsnAsptHrIleMetTyrrAsnIleArgpHeGly 618

1708 AACATCAGTGCTTCACTGAGGAAGTGTATGCAGTGGCAAAATTAGCTGGACTTCATGAT 1766

Db 619 ArgProAspAlaSerAspGluValIleGluAlaAlaLysAlaAlaMetIleHisGlu 638

QY 1768 GCAATCTTCGATGCCATGATATGACACCAAGTAGGGGAAACGAGCTCAAGCTT 1827
 Db |||||
 Db 639 LyslethrSerleupProgluInlathrMetValGlyGluArgGlyLeuLysleu 658
 QY 1828 TCAGAGAGAGAAAGCAAGAGTAGCAATTCAGAGAGCATTTTGAAGGACCCCGAGTC 1887
 Db |||||
 Db 659 SerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 678
 QY 1888 ATACTATGATGAGAGCTACTTCACTGTTAGATTCGATTCGAAAGAGACTTTCTTGGT 1947
 Db |||||
 Db 679 IlePheLeuArgGlyLysIleThrSerAlaLeuMetPheProThrGluArgAlaIleGlnLys 698
 QY 1948 GCCATGAGAGATGTGTGTCACACAGAACTCTTATTTTCATTCGACACAGATGTCAACA 2007
 Db |||||
 Db 699 CysLeuGlyLysLeuCysLysSerArgThrGlyValValAlaIleArgLeuSerThr 718
 QY 2008 GTGGTTGATGAGATGAAATATGTCTTGGATGAGGTAAGTACCGGACGCGTACC 2067
 Db |||||
 Db 719 ValValAlaMetAlaAspLeuIleLeuValLeuAspLysGlyIleIleLeuGluArgGlyAsn 738
 QY 2068 CACCATGGTTTCTTGGTAAACCTCATATGATCTATTCAAGATGTGACATACAGAGC 2127
 Db |||||
 Db 739 HistLysGlyLeuLeuAlaGln---GlnGlyThrIleArgLysSerMetIlePheGluAlaGlnIle 757
 QY 2128 AGCCGTGTGCAAGAACATGATTAACCCCAATGGAGACCAAGAAAGAAATATATCCAAA 2187
 Db |||||
 Db 758 Ala-----GluGlnArgAlaLysSerIleGluLeu 767
 QY 2188 GAGAGGAA 2196
 Db |||||
 Db 768 GlyGluGln 770
 RESULT 15
 F84172
 ABC transport protein [imported] - Halobacterium sp. NRC-1
 C/Species: Halobacterium sp. NRC-1
 C/Date: 02-Feb-2001 #Sequence.Revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: F84172
 R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.
 J.; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabc
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A/Authors: Hou, S.; Daniels, C.J.; Demuth, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
 A/Title: Genome sequence of Halobacterium species NRC-1.
 A/Reference number: A84160; MUID:20504483; PMID:11016950
 A/Accession: F84172
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-642 <STO>
 A/Cross-references: GB:AE004437; NID:G10579770; PIDN:AA818746.1; GSPDB:GNO0138
 C/Genetic:
 A/Gene: trp1
 Alignment Scores:
 Pred. No.: 1.35e-44 Length: 642
 Score: 725.00 Matches: 205
 Percent Similarity: 48.15% Conservative: 120
 Best Local Similarity: 30.37% Mismatches: 250
 Query Match: 17.04% Indels: 100
 Db: 2 Gaps: 17
 AF133659 (1-2345) x F84172 (1-642)
 QY 319 GACCCA---AAAGAGGTTAAAGATGTGATCTCGGAAATCATAAAGCATGCTT 375
 Db |||||
 Db 12 AspProPheGluGluGlnAlaArgAlaAspValAlaAspAlaMetVal-----GlnLeu 28
 QY 376 TCTATGTGTGCCCAAGACAGGCCAGATCTACAGCTAGATGTGCTGCTGCTGGA 435
 Db |||||
 Db 29 PheAspLysIleGlyArgAspHis---SerPheGlnAlaValAlaValAlaValAlaSer 47
 QY 436 TTTTGGGTGTGCAAGGCCAGATATTTGTGTTCCCTTCATGTTAAATATGCTGTA 495
 Db |||||

Db 48 ValPhe-----AlaArgValLeuAspLeuAlaProProValLeuLeuGlyLeuAlaIle 65
 QY 496 GAGAGCCTCAACAGATGTGCGAAAATCATGTAACCTGATGATGACCAATATACAGTT 555
 Db |||||
 Db 66 AspSerValIleGln-----GlyAsnLysAlaPheLeuProPheLeuProGlnSerVal 83
 QY 556 -----GCAACATGGCAAGCAGTT 576
 Db |||||
 Db 84 ValProSerSerLysProAspArgLeuLeuPheMetGlyGlyLeuIleAlaGlySerPhe 103
 QY 577 CTGATGGCTATAGTGTATGATCAAGAGCTGAGCTCTTTTATACCAAGTCCGAAATGCA 636
 Db |||||
 Db 104 LeuLeuGly-----AlaAlaPheHisIleArgPheIleArgAsnIle 115
 QY 637 GTATTGGCAAGTAGCCGAGATTCATCCAGATATACCAAGATATACCAAAATGCTTCTTCAT 696
 Db |||||
 Db 116 GlyPheAsnSerPheSerGlnHisIleGlnHisArgValArgThrAspThrTyArgLys 135
 QY 697 CTTCAACACCTGAGATCTGGTTTCACTGAGCAGACAGACGGGA-----GCT 744
 Db |||||
 Db 136 MetGlnArgLeuAsnMetAspPhePheAlaThrLysGlnThrGlyGluMetSerIle 155
 QY 745 TTAATCTAAGCTATTGACAGAGAAACAGGGGATC-----AGTTT 786
 Db |||||
 Db 156 LeuSerAsnAspValAsnArgLeuGluArgPheLeuAsnAspLysLeuAsnSerAlaPhe 175
 QY 787 GTCCTGAGCTTGGTATTTAATCTTCCATCATGTTGAAGTATGATGCTTGCAGT 846
 Db |||||
 Db 176 ArgLeuSerValMetVal-----LeuAlaIle 184
 QY 847 GGTGTTTGTATTACAATGCGGTGCCAGTTGCTTGGTAACTTGGACACTTGGT 906
 Db |||||
 Db 185 GlyValTyLeuPheValValAsnTrpGlnLeuAlaValLeuThrMetLeuProValPro 204
 QY 907 ACATACACAGATTCACAGTGCACACAGCTGAGAGACTGATTTAGATTAAGAATG 966
 Db |||||
 Db 205 IleIleAlaLeuPheThr-----TyrArgPheValAlaMetIleGlnPro 219
 QY 967 AACAAAGCATATGATGATGAGGATGATGCTGATGACTGACTGCTG----- 1014
 Db |||||
 Db 220 LysTyAlaAspValArgSerSerValGlyHisLeuAsnSerArgLeuGluAsnAsnLeu 239
 QY 1015 ---AATTATGAACCTGATGATTTTAAATGAAGATATGAAACACAGATATGAT 1071
 Db |||||
 Db 240 GlyGlyIleGlnValIleLysThrSerAsnThrGluArgTyGluSerAspArgValAsp 259
 QY 1072 GGAATTTTGAAGACGATGAGACTGCTTCAATGAAGATCACTGCTGCTATGCTG 1131
 Db |||||
 Db 260 AspValSerGlnGlyTyArgPheAspAla----- 268
 QY 1132 AACTTGGTCAAGTGTCTATTTTCACTGTCGTTTAAACGTATATATGTCGCGCAGT 1191
 Db |||||
 Db 269 AsnTrpGlyAlaIleThrIleArgIleLysPhePheProAlaLeuArgIleIleSerGly 288
 QY 1192 CAGGGAATTGTG----- 1203
 Db |||||
 Db 289 ValGlyPheValLeuThrPheValIleGlyGlyIleTrpValAlaThrArgLysProPheMet 308
 QY 1204 -----GACAGTACCTTACCTGTTGAGATCTAGTAATGGTGAATGACTGCTTTTCA 1257
 Db |||||
 Db 309 PhePheSerGlyThrLeuAspProGlyGluPheValThrPheIleLeuLeuSerGlnGln 328
 QY 1258 CTTTCAATACCCCTGAACCTTCTGGAACGTATATAGAGAGCTAGACAGACTCAT 1317
 Db |||||
 Db 329 PheIleTrpProMetAlaGlnPheGlyGlnIleIleAsnMetTyGlnArgAlaArgAla 348
 QY 1318 GATATGACACCTTGTATTCTACTCAAGGTAGACACCAAAATTAAGCAAAAGTATG 1377
 Db |||||
 Db 349 SerSerGluArgIlePheGlyLeuMetAsnGluProSerArgIleGluGluAsnProAsp 368
 QY 1378 GCATCTCCCTTCAGATCACACACAGACAGACGCTACCGTCCCTTGTATATGCTGATTT 1437
 Db |||||
 Db 369 AlaAspGluLeuValVal-----AspArgGlyGlyValValTyArgAspAspValArgPhe 386

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 05:55:00 ; Search time 17.5 Seconds
(without alignments)

6847.099 Million cell updates/sec

Title: AF133659
Perfect score: 4254
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 140259 segs, 2554876 residues

Total number of hits satisfying chosen parameters: 280518

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=x1h
-Q=/cg2_1/USPTO.spool/AF133659/runat_10022003_155442_29603/app_query.fasta_1.2503
-DB=Published.Applications_AA -QMT=fastan -SUFIX=rapb -MINMATCH=0.1
-LOOPCH=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=AF133659 @CGN 1.1 10 @runat_10022003_155442_29603
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGQUERY -NEG SCORES=0 -NAMES -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELext=7

Database : Published Applications AA:*

- 1: /cg2_6/prodata/2/pubpaa/US08 NEW PUB pep:*
- 2: /cg2_6/prodata/2/pubpaa/PCT_NEW PUB pep:*
- 3: /cg2_6/prodata/2/pubpaa/US06 NEW PUB pep:*
- 4: /cg2_6/prodata/2/pubpaa/US07 PUBCOMB pep:*
- 5: /cg2_6/prodata/2/pubpaa/US07 NEW PUB pep:*
- 6: /cg2_6/prodata/2/pubpaa/US07 PUBCOMB pep:*
- 7: /cg2_6/prodata/2/pubpaa/PCTUS PUBCOMB pep:*
- 8: /cg2_6/prodata/2/pubpaa/US08 PUBCOMB pep:*
- 9: /cg2_6/prodata/2/pubpaa/US09 NEW PUB pep:*
- 10: /cg2_6/prodata/2/pubpaa/US09 PUBCOMB pep:*
- 11: /cg2_6/prodata/2/pubpaa/US10 NEW PUB pep:*
- 12: /cg2_6/prodata/2/pubpaa/US10 PUBCOMB pep:*
- 13: /cg2_6/prodata/2/pubpaa/US60 NEW PUB pep:*
- 14: /cg2_6/prodata/2/pubpaa/US60 PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1247	29.3	836	10	US-09-953-688A-7
2	1201.5	28.2	475	9	US-10-076-157-4
3	1183	27.8	574	10	US-09-953-688A-1
4	736.5	17.3	640	9	US-09-976-059-9

5	699.5	16.4	578	10	US-09-815-242-12501	Sequence 12501, A
6	685	16.1	621	9	US-09-738-626-6020	Sequence 6020, Ap
7	685	16.1	621	10	US-09-953-253-2	Sequence 2, Appli
8	654	15.4	582	10	US-09-815-242-14006	Sequence 14006, A
9	651	15.3	582	10	US-09-815-242-10115	Sequence 10115, A
10	632	14.9	587	9	US-10-260-877-26	Sequence 26, Appl
11	632	14.9	587	10	US-09-815-242-10960	Sequence 10960, A
12	618	14.5	603	10	US-09-815-242-5202	Sequence 5202, Ap
13	615.5	14.5	1275	10	US-09-749-340-6	Sequence 6, Appli
14	593.5	14.0	1247	9	US-09-738-626-4751	Sequence 4751, Ap
15	593	13.9	1195	10	US-09-873-409-6	Sequence 6, Appli
16	590.5	13.9	571	10	US-09-815-242-10619	Sequence 10619, A
17	590	13.9	760	10	US-09-833-017-26	Sequence 26, Appl
18	585	13.8	1276	10	US-09-866-866A-6	Sequence 6, Appli
19	582.5	13.7	766	9	US-10-072-621-6	Sequence 6, Appli
20	582.5	13.7	766	10	US-09-795-693-17	Sequence 17, Appl
21	578.5	13.6	1280	10	US-09-866-866A-4	Sequence 4, Appli
22	575.5	13.5	1394	9	US-10-101-388-3	Sequence 3, Appli
23	573.5	13.5	1276	10	US-09-866-866A-8	Sequence 8, Appli
24	572.5	13.5	1280	9	US-10-072-621-7	Sequence 7, Appli
25	572.5	13.5	1280	10	US-09-866-866A-2	Sequence 2, Appli
26	569.5	13.4	1222	10	US-09-873-409-5	Sequence 5, Appli
27	567	13.3	1272	10	US-09-769-097-4	Sequence 4, Appli
28	566	13.3	1272	10	US-09-769-097-2	Sequence 2, Appli
29	565.5	13.3	583	10	US-09-815-242-13458	Sequence 13458, A
30	562.5	13.2	1280	9	US-09-758-828-2	Sequence 2, Appli
31	562	13.2	1280	9	US-10-044-671-2	Sequence 2, Appli
32	554.5	13.0	656	9	US-09-738-626-4577	Sequence 4577, Ap
33	535.5	12.6	659	10	US-09-873-409-1	Sequence 1, Appli
34	535.5	12.6	812	10	US-09-873-409-2	Sequence 2, Appli
35	535.5	12.6	1058	10	US-09-873-409-4	Sequence 4, Appli
36	528.5	12.4	748	9	US-09-870-759-41	Sequence 41, Appl
37	513	12.1	400	10	US-09-765-272-190	Sequence 190, App
38	509	12.0	1263	9	US-09-882-694-11	Sequence 11, Appl
39	485	11.4	577	9	US-09-738-626-4578	Sequence 4578, Ap
40	472.5	11.1	514	10	US-09-873-409-8	Sequence 8, Appli
41	470.5	11.1	646	10	US-09-841-133-567	Sequence 567, App
42	464.5	10.9	659	10	US-09-841-133-497	Sequence 497, App
43	461.5	10.8	551	10	US-09-815-242-11394	Sequence 11394, A
44	449	10.6	541	10	US-09-873-409-7	Sequence 7, Appli
45	447.5	10.5	339	9	US-09-764-884-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-953-688A-7
; Sequence 7, Application US/09953688A
; Patent No. US20020102649A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Gorgone, Gina
; APPLICANT: Corley, Neil C.
; APPLICANT: Aizimai, Yaida
; APPLICANT: Patterson, Chandra
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN TRANSPORT PROTEIN HOMOLOGS
; FILE REFERENCE: PF-0555 US
; CURRENT APPLICATION NUMBER: US/09/953,688A
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/113,427
; PRIOR FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 836
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
; FEATURE:
; OTHER INFORMATION: 2982567, Genbank
US-09-953-688A-7

Alignment Scores:

Pred. No.:	7.54e-104	Length:	836
Score:	1247.00	Matches:	280
Percent Similarity:	54.94%	Conservative:	126
Best Local Similarity:	37.89%	Mismatches:	247
Query Match:	29.31%	Indels:	86
DB:	10	Gaps:	13

AF133659 (1-2345) X US-09-953-688A-7 (1-836)

Oy	28	CGCTGGGCGCCGCGGGCGGTCTTTGCAGAAAGCGCGGCACTCCGCGATTCTGATC---	84
Db	133	ArgGlnSerLeuAlaMetGlyValTrpMetPheArgHisSerLeuGlyLeuLeu	1522
Oy	85	-----CGGCTTACTTCCTGTACCGGCTCAGGTCGGCAG	120
Db	153	LeuTrpThrValThrPheAlaAlaGlnIleuLeuValLeuValSerTPasMetSerProGln	1722
Oy	121	TGGAGGCCAACAACCTCGGCGCCTTGGGAACCCCTCGAGCTCACGATTCCAGAGTCA	180
Db	173	Trp-----	173
Oy	181	TTTTAAAGTATCACATGCGAGAGATTGGAAAAGSCAATTCAGACACTTTAGATGCT	240
Db	174	-----TrrTrpSerArgLalaaspLeuGlnGlnValGlnPhe-	186
Oy	241	GCAAAGGCTTCACGATrTGCCCACTGATATAAAAAAGAGACA--	282
Db	187	-----GlyLeuTrpValLeuArgTyrlMetThrSerGlyGlyLeuPheIleu	202
Oy	283	--TGTTGG--	288
Db	203	GlyLeuTrpAlaProGlyLeuArgProGlnSerTyrlThrIleuHisValAsnGluAsp	222
Oy	289	---CATGGTCATGCGAGGAGGAGACTCCACACAGACCMAAGAAGGTTAAAGATGT	345
Db	223	GlnaspGlyGlyArgAsnGlnGlyArgSerThrAspProAsnSerThrTrpArgAspLeu	242
Oy	346	GATACTCGAAAATCATATAAAGCAATGCTTTTATGTGTGGCCCCAAGAAGACAGCCAGAT	405
Db	243	GlyArgLys-----LeuArgLeuIleuSerGlyTyrlLeuTrpProArgLysProSer	260
Oy	406	CTACAGAGCTAGAGTTGCCATTTCGTGGGATTTTGGTSGTGCMAAGCCATGAAATAT	455
Db	261	LeuGlnLeuThrValLeuLeuIleuCysMetGlyLeuMetGlyLeuAspArgAlaLeuAsnVal	280
Oy	466	GNGGTTCCTCATGTTTAAATATATCTGTACAGCGCTCAACCGAGATGTCGGAAACATG	525
Db	281	LeuValProIlePheTherTyraArgAspIleValAsnLeuLeuThr-----	294
Oy	526	CTGAACCTGAGTAGATGACCA-----AATACAGTTGCACAGGACAGACAGTTCTG	579
Db	295	-----SerLysAlaProTrpSerSerIleuAlaTrpThrValThrTrpValPhe	311
Oy	580	ATTGGCTAT-----GGTGTATCAAGAGCTGAGAGCTGCTTTTTTTAACGAATTCGAAT	633
Db	312	LeuLysPheLeuGlnGlnGlyGlyThrGlySerThrGlyPheValSerAsnLeuArgThr	331
Oy	634	GCAGATTTTGGCAAGGTGACCCCAAGATTCATCCGAGAATATAGCCAATAATGCTTCTC	693
Db	332	PheLeuTrpIleArgValGlnGlnPheThrSerArgGlyValGluLeuArgLeuPheSer	351
Oy	694	CATCTTCAACAACCTGATCTGGGTATTTACCTGTGACAGACAGACGGAGCTTATCAAG	753
Db	352	HisLeuHisGluLeuSerLeuArgTrpHisIleuGlyArgGlnThrGlyValLeuArg	371
Oy	754	GCTATTGACAGAGAACAGGGGATACGATTGTCTGACTGCTTGTGATTTAATCTT	813
Db	372	IleValAspArgGlyThrSerSerValThrGlyLeuIleuSerTyrlLeuValPheAsnIle	391
Oy	814	CTTCCATCATGTTTGAAGTGAATGCTGTCTGAGTGTTTGTATATCAAAATGC----	867

[illegible]

```

QY 1948 GCCATGAAAGATGTCGTCACAAACAGAACTTCTATTCTTTCATTCGACACAGATTGTCAACA 2007
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 SerLeuAlaLysValCysThrAsnArgThrPheIleValAlaIleHisArgLeuSerThr 787

QY 2008 GTGTGATGAGATGAAATCATTTGTGATCATGAGGTAAAGGTAAGCCGAAACGNGATACC 2067
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 788 ValValAsnAlaAspGlnIleLeuValIleLysAspGlyCysIleIleGluArgGlyArg 807

QY 2068 CACCATGGTTTGCTTGCTTAACCTCATAGTACTTATTCAGAAATGTGGCATACACAG 2124
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 808 HisGluAlaLeuLeuSerArg---GlyGlyValTyrAlaGluMetIlePheGlnLeuGln 825

RESULT 2
US-10-076-157-4
; Sequence 4, Application US/10076157
; Publication No. US20030027309A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Suelberger, Harald
; APPLICANT: Josefken, Hans Wolfgang
; APPLICANT: Doval, Jose Luis Revuelta
; APPLICANT: Jimenez, Alberto;
; TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii and the use the
; TITLE OF INVENTION: in
; FILE REFERENCE: 48684DIV
; CURRENT APPLICATION NUMBER: US/10/076,157
; PRIOR APPLICATION NUMBER: 2002-02-15
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: WordPerfect v. 6.1
; SEQ ID NO 4
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Ashbya gossypii
; US-10-076-157-4

Alignment Scores:
Pred. No.: 7 37e-100 Length: 475
Score: 1201.50 Matches: 234
Percent Similarity: 71.74% Conservative: 96
Best Local Similarity: 50.87% Mismatches: 127
Query Match: 28.24% Indels: 3
DB: 9 Gaps: 2

AF133659 (1-2345) x US-10-076-157-4 (1-475)

QY 757 ATTGACAGAGAACAGAGGGATGATCTTTGTCCTGAGTGGTCTTGATTAATCTTCTT 816
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MetAspArgGlyCysIleSerTyrValLeuSerAlaMetValPheHisIleIle 20

QY 817 CCCATCATGTTTGAAGTATGCTTGTGATGCTGTTTGTATTAACAATGCGGTCCAG 876
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 ProIleThrPheGlnIleSerMetValCysGlyIleLeuThrTyrGlnPheGlyAlaSer 40

QY 877 TTTCCTTTGGTAACCTTGAACACTTGTGATACATACACAGATTCACAGTTGCAGTACA 936
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 PheIleAlaIleThrPheSerThrMetLeuLeuTyrSerIlePheThrPheArgThr 60

QY 937 CGGTGAGAACTAGATTAGAAATGAATGAACAACAGATTAATGACAGGTATGCT 996
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 AlaTyrArgThrArgPheArgArgAspAlaAsnLysAlaAspAsnLysAlaAlaSerVal 80

QY 997 GCTATGACTCACTGCTGAATATGAACCTGTGAAGTATTTTAATTAAGAAAGATGAA 1056
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 81 AlaLeuAspSerLeuIleAsnPheGluAlaValLysTyrPheAsnAsnGluLysTyrLeu 100

QY 1057 GCACAGATATGATGATTTTGAACGATACAGATCTCTTCATTTGAAAGATCCTCT 1116
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 101 AlaAspLysTyrHisThrSerLeuMetLysTyrArgAspSerGlnIleLysValSerGln 120

```

```

QY 1117 ACTGTGCTATGCTGAACCTTGGTCAAAAGTCTATTTCAGTGTGCTTAAACGATA 1176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 SerLeuAlaPheLeuSerThrGlyGlnAsnLeuIlePheThrThrAlaLeuThrAlaMet 140

QY 1177 ATGTGTGCTCCAGTCAGGGAATTTGGCAGGTACCTTATCTGTGGAGTCTAGTAAAG 1236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 MetTyrMetAlaCysAsnGlyValMetGlnGlySerLeuThrValGlyAspLeuValLeu 160

QY 1237 GTAATGAGATGCTTTTTCAGTTTCACTTATACCCCTGAACTTTGTGGAACTGATATAGA 1296
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 IleAsnGlnLeuValPheGlnLeuSerValProLeuAsnPheLeuGlnLysSerValTyrArg 180

QY 1297 GAGACTAGAACACACACTCATATGATATGAACACCTGTTTACTACTCAAGTAGACACC 1356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 AspLeuLysGlnSerLeuIleAspMetGlnSerLeuPheLysLeuGlnLysAsnGlnVal 200

QY 1357 CAATTTAAAGACAAAGATGATGATCTCCCTTCAGATCACACACAGACGTAACCGTG 1416
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 ThrIleLysAsnSerProAsnAlaGlnAsnLeuProIle---HisLysProLeuAspIle 219

QY 1417 GCCTTTGATATGATGATTTGAAATCATTTAGAGGCCAGAAAGTCTTAGTGAATATCC 1476
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 ArgPheGlnAsnValThrPheGlyTyrAspProGluArgArgIleLeuAsnAsnValSer 239

QY 1477 TTGGAAGTCCCTGACAGAAAGAAAGTGCCTTGTAGAGGTGGGTGCGAGAAAGC 1536
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 PheThrIleProAlaGlyMetLysThrAlaIleValGlyProSerGlySerGlyLysSer 259

QY 1537 ACAATAGTGAAGCTATATTATTCCTTCTATGAGCCTTCAAAAGGTAGCATTTATCTTGCT 1596
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 ThrIleLeuLysLeuValPheArgPheTyrGluProGlnGlnIleLysIleValGly 279

QY 1597 GGTCAAAATATACAGATGAGCTGGAAGCCTTGGAGGCGAGTGGAGTGTACT 1656
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 GlyThrAspIleArgAspLeuAspLeuSerLeuArgLysAlaIleGlyValPro 299

QY 1657 CAGATGCTGTCTCTTCCATATATCTATTATTAACAACCTTTATATGAAACATCAGT 1716
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 GlnAspThrProLeuPheAsnAspThrIleTyrGlnAsnValLysPheGlyAsnIleSer 319

QY 1717 GCTTACCTGAGGAAGTATGATGATGCGCAAAATTTGCTGCACTTATGATGCAATTTCT 1776
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 SerSerAspAspIleIleLeuArgAlaIleGlyLysAlaGlnLeuThrLysLeuGln 339

QY 1777 CGAATGCCATGATATATGACACCAAGTAGGGGAGGAGACTCAAGCTTTGAGAGGA 1836
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 AsnLeuProLysGlyAlaSerThrValValGlyLysGlyLeuMetIleSerGlyGly 359

QY 1837 GAAACCAAAAGATAGCAATTTGCAAGAGCATTTTGAAGAGCCCCAGCATACTCTAT 1896
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 GlnLysGlnArgLeuAlaIleAlaArgValLeuLeuLysAspAlaProLeuMetPhePhe 379

QY 1897 GATGAAGTACTTTCATGTTAGTTGATTAAGTGAAGACATATTTCTGTGTCATGAAG 1956
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 AspGluAlaThrSerAlaLeuAspThrHisThrGlnGlnAlaLeuLeuHisThrIleGln 399

QY 1957 GAT-----GTGTCAAAACAGAACTTCTATTATTTGACACAGATTTGTAACAGTG 2010
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 GlnAsnPheSerSerAsnSerLysThrSerValTyrValAlaHisArgLeuAspGlyThrIle 419

QY 2011 GTTGATGAGATGATGATCAATTTGCTTGATGAGGGTAAAGTGAAGCCAGTGTGATCCAC 2070
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 AlaAspAlaAspLysIleIleValLeuGlnGlnGlySerValArgGlnGlnGlyThrHis 439

QY 2071 CATGTTTCTGCTGTAACCTCATATGATATCTTCAAGAAATGTGSCATACACAGACACC 2130
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 SerSerLeuLeuAlaSerGlnGlnSerLeuTyrArgGlyLeuThrAspIleGlnGlnAsn 459

RESULT 3
US-09-953-688A-1
; Sequence 1, Application US/09953688A
; Patent No. US20020102649A1

```

```

: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Yue, Henry
: APPLICANT: Reddy, Roopa
: APPLICANT: Gorgone, Gina
: APPLICANT: Corley, Neil C.
: APPLICANT: Azimzai, Yalda
: APPLICANT: Patterson, Chandra
: APPLICANT: Baughn, Mariah R.
: TITLE OF INVENTION: HUMAN TRANSPORT PROTEIN HOMOLOGS
: FILE REFERENCE: PP-0555 US
: CURRENT APPLICATION NUMBER: US/09/953,688A
: PRIOR FILING DATE: 2001-09-12
: PRIOR APPLICATION NUMBER: 09/113,427
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 574
: TYPE: PRT
: ORGANISM: HOMO SAPIENS
: FEATURE:
: OTHER INFORMATION: 2074412, ISLNOT01
: US-09-953-688A-1

Alignment Scores:
Pred. No.: 3,686-98 Length: 574
Score: 1183.00 Matches: 250
Percent Similarity: 62.12% Conservative: 114
Best Local Similarity: 42.66% Mismatches: 194
Query Match: 27.81% Gaps: 28
DB: 10

AF133659 (1-2345) x US-09-953-688A-1 (1-574)

QY 430 CTGGATTTTGGGTGGTGCAAGGCCATGATATTGGTTCCTTCATGTTAAATAT 489
DB 1 MetGlyLeuMetGlyValArgArgAlaLeuAsnValPheValProIlePheTyrArg-- 19
QY 480 GGTGTAGACAGCTCAACCAAGATGTCCGGAACATGCTGACCTG-----AGTATGCA 543
DB 20 -----AsnIleValAsnLeuLeuThrGluLysAla 29
QY 544 CCA---AATACAGTTGCAACCATGACACAGCA-----GTTCTGATTGGC 585
DB 30 ProThrPheSerLeuAlaThrPheValThrSerTyrValPheLeuLysPheLeuGlnGly 49
QY 586 TATGGTGTATCAAGAGCTGAGCTGCTTTTAAAGAAATGCAAGTCAATGATTTGGC 645
DB 50 GlyGlyThrGlySerThrGly-----PheValSerAsnLeuArgThrPheLeuTrpIle 67
QY 646 AAGTAGCCCAAGATTCATCCGAAGATAGCCAAAATGCTTTCTCTCATCTTCACAAAC 705
DB 68 ArgValGlnGlnPheThrSerArgArgValGluLeuLeuIlePheSerHisLeuHisGlu 87
QY 706 CTGGATCTGGTTTCACTGTGAGCAGACAGAGGGAGCTTATCTTAAGCTATTGACAGA 765
DB 88 LeuSerLeuArgTrpHisLeuGlnArgArgGlnGlyGluValLeuArgLysAlaAspArg 107
QY 766 GGAACAAGGGATATCATGTTTGTCTGAGTGGTGTGTTGTTAATCTTCTCCATCATG 825
DB 108 GlyThrSerSerValThrGlyLeuLeuSerTyrLeuValPheAsnValIleProThrLeu 127
QY 826 TTTGAAGTAGATCTGTGCTGAGTGTGTTGTTGTTATACAAATGC-----GTCGCCACGTTT 879
DB 128 AlaAspIleIleIleIle-----GlyIleIleTyrPheSerMetPhePheAsnIleTrpPhe 145
QY 880 GCTTTGGTAACTCTTGAACACTGGTGTATCATACAGATTCACAGTTGAGTTCAGTACAGG 939
DB 146 GlyLeuIleValPheLeuLysMetSerLeuTyrLeuThrIleValValThrGlu 165
QY 940 TGGAGAACTAGATTGATAGAAAAGAAAGCAAGAGATTAATGATGACAGTAACTCTGCT 999
DB 166 TrpArgThrLysPheArgAlaMetAsnThrGlnLysAsnAlaThrArgAlaArgAla 185
QY 1000 ATAGACTCAGCTGCAATTAAGAACTGTGAGATTTTAAATGAAGATATGAGCA 1059
DB 186 ValAspSerLeuLeuAsnGlnThrValLysTyrTyrAsnAlaGluSerTyrGluVal 205
QY 1060 CAGAGATATGATGATTTTGAAGACGATGAGACTGCTTCATTTGAAAAGTACCTTACT 1119
DB 206 GluArgTyrArgGluAlaIleIleLysTyrGlnLysLeuGluTrpLysSerSerAlaSer 225
QY 1120 CTGGCTATGCTGAACCTTTGGTCAAGTGTCTATTTCAGTGTGCGTTTAAACGCTATATG 1179
DB 226 LeuValLeuLeuAsnGlnThrGlnAsnLeuValIleGlyLeuGlyLeuAlaGlySer 245
QY 1180 GTGCTCGCCAGCTGACGGGATATGTGGCAGGTACCTTACTGTGTGAGATCTAGTATGTTG 1239
DB 246 LeuLeuCysAlaTyrPheValThrGlnGlnLysLeuGlnValGlyAspTyrValLeuPhe 265
QY 1240 AATGACTGCTTTTTCAGCTTTTCATTAACCCCTGAACTTTCTGGGAAGCTGATATAGAG 1299
DB 266 GlyThrTyrIleIleGlnLeuTyrMetProLeuAsnTrpPheGlyThrTyrTyrArgMet 285
QY 1300 ACTAGACAGACCTCATATGATATGACACACTTGTACTCTACTCAAGTAGACACCCAA 1359
DB 286 IleGlnThrAsnPheIleAspMetGluAsnMetPheAspLeuLeuLysGluGluThrGlu 305
QY 1360 ATTAAGACAAGAGATGGCATCTCCCTTCAGATCAACACAGACAGATACCGTGCC 1419
DB 306 ValLysAspLeuProGlyAlaGlyProLeuArgPhe-----GlnLysGlyArgIleGlu 323
QY 1420 TTTGATATGTGCTTTTGAATATACATTAGAGGGCCAGAAAGCTTTAGTAAATTCCTTT 1479
DB 324 PheGluAsnValHisPheSerTyrAlaAspLysArgGluThrLeuGlnAspValSerPhe 343
QY 1480 GAATCCCTGCAAGAAAGAAAGTGGCCATTGTAGAGTAGTGGTCAAGGAAAGACACA 1539
DB 344 ThrValMetProGlyGlnThrLeuAlaLeuValGlyProSerGlyAlaGlyLysSerThr 363
QY 1540 ATAGTAGGCTTATTTTGGCTTCTATGAGCTCAAAAGGTACATTTATCTTGCTGGT 1599
DB 364 IleLeuArgLeuLeuPheArgPheTyrAspIleSerSerGlyCysIleArgIleAspGly 383
QY 1600 CAAATATACAGATGTGAGCTGGAAGCCCTTGAGAGGGCAGGAGTGTGATCTCAG 1659
DB 384 GlnAspIleSerGlnValThrGlnAlaSerLeuArgSerHisIleGlyValValProGln 403
QY 1660 GATGCTGCTCTTCATATATCTATTATTATTAACAACCTTTATATGGAACATAGTGGCT 1719
DB 404 AspThrValLeuPheAsnAspThrIleAlaAspAsnIleArgTyrGlyArgValThrAla 423
QY 1720 TCACCTGAGAGGTATGACGTGGCAAAATTACCTGAGCTTCATGATGCAATCTCGA 1779
DB 424 GlyAsnAspGluValGluAlaAlaAlaGlnAlaAlaGlyIleHisAspAlaIleMetAla 443
QY 1780 ATGCCACATGATATGACACCAAGTAGGGGAGAGAGACTCAAGCTTTCAGAGAGAA 1839
DB 444 PheProGluGlyTyrArgThrGlnValGlyGluArgGlyLeuLysSerGlyGlyGlu 463
QY 1840 AAGCAAGAGTAGCAATTTGCAAGACCAATTTGAAGACCCCAAGTCATATCTATGAT 1899
DB 464 LysGlnArgValAlaIleAlaIleArgThrIleLeuLysAlaProGlyIleIleLeuLeuAsp 483
QY 1900 GAAGCTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1959
DB 484 GluAlaThrSerAlaLeuAspThrSerAsnGlnArgAlaIleGlnAlaSerLeuAlaLys 503
QY 1960 GTTGCTAAACACAGAACTTATTTTATTCATGACACAGATTTGCAAGCTGTGATGCA 2019
DB 504 ValCysAlaAsnArgThrThrIleValValAlaHisArgLeuSerThrValValAsnAla 523
QY 2020 GATGAATATGTTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2079
DB 524 AspGlnIleLeuValIleLysAspGlyCysIleValGluArgGluArgHisGluAlaLeu 543

```


Db 480 ileavgaiaasleuleuylrgrproaspalaiglygluargaspulevalialaia 499
Qy 1744 GCAAAATTAAGCTGAGCTTATGATGCAATTTCTCCAAATCCACATGATATACACCCAA 1803
Db 500 CysgluaialaiaargliletrpgleuvaliserseuProaspdlYleuasprrval 519
Qy 1804 GAGAGGAAACAGAGCTTCAAGCTTTCAGAGAGGAGAAAGCAAGATGCAATTCAGAGA 1863
Db 520 valigvayasparglylYrargleuaserglygluluglinalrgleuhalaleuhalarg 539
Qy 1864 GCCATTTTGAAGACCCCTCACTACTCTATGATGAAGACTTCAATTCATGTTAGATTGG 1923
Db 540 leuleuYserProProvalValValleuaspglualatThrAlahisLeuaspser 559
Qy 1924 ATTACTGAAGACTATTTCTTGCTCCAGAGAGATGTCGTAACACAGACTTCTATT 1983
Db 560 gluserglualialaleginalrgalaleuasprrhalaleuhalaglyarghrserleu 579
Qy 1984 TTCATTTGCACAGATTCACAGCTGTGATGCAAGATGATGAATGATGTTGTGATCG 2043
Db 580 valileahiahirgleuhalatThrileuaspalaaspghilleuvalilleaspasp 599
Qy 2044 GGTAAAGTACCCGAACTGTGACCCAGCCAGTGTGCTTGAACCTTCATAGATCTAT 2103
Db 600 glYargvalialgluarglylYrhrhsaerghileuileahis--glygluleuYr 618
Qy 2104 TCAGAAATGTGGCATACACAG 2124
Db 619 AlaGluleuYrarghrghin 625

RESULT 5

US-09-815-242-12501

Sequence 12501, Application US/09815242

Patent No. US2002061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zykien, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815.242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12501

LENGTH: 578

TYPE: PR1

ORGANISM: Staphylococcus aureus

US-09-815-242-12501

Alignment Scores: 1.13e-54 Length: 578

Score: 699.50 Matches: 181
Percent Similarity: 51.93% Conservative: 128
Best Local Similarity: 30.42% Mismatches: 263
Query Match: 16.44% Indels: 23
DB: 10 Gaps: 11

AF133659 (1-2345) x US-09-815-242-12501 (1-578)

Qy 358 ATTCATTAAGAGATGCTTTCTTATGTGGCCAAAGACAGAGCCAGATCTACAGCTAGA 417
Db 1 MetileuysarglyrleuglnphevallyeProlyr-----lyrYrarglilephe 17
Qy 418 GTTGCAATTTGCTGGAGATTTTGGGTGGCAAGCCATGAATATTGTTCCCTTC 477
Db 18 Alathrillevallelglyleile-----lyrphleglylePrometeuileProleu 35
Qy 478 ATGTTTAATTTGCTGTGACAGCCTCAACAGATGTCCGAAACATGCTGAACCTGAGT 537
Db 36 leuileuYrYralaialeaspglyval-----lleasmenhlsialeuThrThrAsp 53
Qy 538 GATGACCAAAATPACAGTTGCAACCATGGCAACAGACAGATGCTGATGGTATGATCA 597
Db 54 GlulysvalihshisleuThrillealalileglyllealaleuPheillephevalilleval 73
Qy 598 AGAGCTGAGCTGCTTTTTCAGAAATTCGAAATGCAATGCAATGTTGGCAAGTACCCAG 657
Db 74 ArgProProillegluphe-----llearglnYrleuhalaglnYrThrSerAsn 90
Qy 658 AATTCATCCGAAAGATGCCAAATAATGCTTTTCCATCTTCACAACTGATGCTGGGT 717
Db 91 lysileuYrYrAspillearglylsleuYrAsnhsleuGlnalaleuSerAlaarg 110
Qy 718 TTTACCTGACAGACAGACAGAGGGA---GCTTATCTTAAGCTATTGACAGAGAAACAG 774
Db 111 PherYrAlaasnaanglnvalglylnvalileserYrvalilleasmpvalgluGln 130
Qy 775 GGTATCAGTTTGTCTGAGTGTGTTGGTATTAATCTTCTCCATCATGTTGAAATG 834
Db 131 ThrlysaPheilleuThrGlyleuMet---AsnileYrleuaspCyileThrile 149
Qy 835 ATGCTGTGACGTGTTTGTATTAACAAATGCGGTGCCAGTTTGCTTGGTAACTT 894
Db 150 llelealaleuSerileMetPhePhe-----leuaspVallyleuThrleu 165
Qy 895 GGAACACTGTGATACATCACAGCATTCAGCTGACGACAGCGTGAGAAATGATTT 954
Db 166 AlaaleuPheillepheProPheYrilleuThrYalYrValphePheglyArgleu 185
Qy 955 AGA-----ATAGAATGACAAAGCAGATATGATGACAGGTAACTGCTATAGAC 1005
Db 186 ArglylsleuThrarggluargserglnalaleuhalaglnvalinglylPheleuHlgly 205
Qy 1006 TCACGTGCTGAATTTGAACCTGTGAATTTTAATATGAATAATATGAAGCAGAGA 1065
Db 206 ArgvalinglnYlleservalVallyserPhehalaleglnuaspasglualalyasn 225
Qy 1066 TATGATGATTTTGAAGACGACGCTCATGTAAGAAATGACCTACTCTGCT 1125
Db 226 PheaspYlsYsaenThrAsnphleuThrarglialeuYshslnArgYrtpasnha 245
Qy 1126 ATGCTGAATTTGGTCAAGTCTATTTCAGTGTGGTTCATGACGATATAGTGCTC 1185
Db 246 TyserPhealalealeasnthrValthrAspilleglyProillevalilleglyVal 265
Qy 1186 GCCAGTCAGGAATTTGCGCAGGTACCTTACCTTGGAGATCTAGATATGAGTGAATGA 1245
Db 266 GlylatYrleuAlaialeserglyserilePheYrValglyThrleuAlaalephevalgly 285
Qy 1246 CTGCTTTTCACTTTTATTCCTCCGTGAATCTTGGGAATGTATATAGAGACTAGA 1305
Db 286 TyrleuGlnleuPheglyProleuYrargYrleuValalaserPheThrThrleuThr 305
Qy 1306 CAAGCACTCATATGATGAACACCTGTTTACTACTCAAGTATGACACCAATTTAA 1365


```

Db 306 GlnSerPheAlaSerMetLeuAspArgValAlaPheGlnLeuIleLeuPoluAspIlyrThrSpleu 325
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 1366 GACAAAGTATGATGCATATCCCCCTTCAGATACACCAAGACAGTACCGTGCCCTTTGAT 1425
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 326 AsnGlyValGlyIaGlnProIleGluIle-----LysGlnIlyrGlyLeaSprIleAsp 343
Qy 1426 AATGTGCTTTTGAATACATTGAGGGCCAG---AAAGTCCCTTAGTGAATATCCCTTGA 1487
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 344 HisValSerPheGlnIlyrAsnAspArgGlnAlaProIleLeuIleAspIleLeuSer 363
Qy 1483 GTCCCTCAGAGAAAGAAAGTGGCCATTTGTAGAGAGTATGTGGGTCAAGGGGAAACACAAATG 1542
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 364 IlegIuylsGlyuIthrValAlaPheValGlyMetSerGlyGlyIylsSerThrIleu 383
Qy 1543 GTGAGCTATTATTTGCTTCCTTCATGAGCCCTCAAAAGGGTAGCATTTATCTTGCTGTCAA 1602
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 384 IleAsnLeuIleProGlyrPheIlyrAspValThrSerGlyGlnIleLeuIleAspGlyHis 403
Qy 1603 AATATACAAAGTGTGACCCCTTGAAAGCCTTCGGAGGGCAGTGGAGTGGTACTACAGAT 1662
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 404 AsnIleIlyrAspPheLeuThrGlySerLeuAlaArgGlnIleGlyLeuValAlaGlnAsp 423
Qy 1663 GCTGTCCCTTCCTCAATATACATATTATTAACAACCTGTATATAGAAACATCAAGTGCCTCA 1722
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 424 AsnIleLeuPheSerAspThrValIylsGlyAsnIleLeuLeuIlyrGyrProThrAlaThr 443
Qy 1723 CTTGAGAAAGTGTATGCAAGTGGCAAAATTAAGCTTGACCTTCATGATCAATCTTCGAATG 1782
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 444 AspGlnGluValValGluAlaIleAlaIylsMetAlaAsnAlaHisAspPheIleMetAsnIleu 463
Qy 1783 CCACATGATATGTGACACCCCAAGTATGGGGAAACAGAGACTCAAGCTTTCAGAGAGAGAAAAG 1842
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 464 ProGlnGlyIlyrAspThrGlnValGlyGluAlaGlyValIylsLeuSerIylGlyGlnIyls 483
Qy 1843 CAAGAAGTAGAATTTGCAAGAGCCATTTTGAAGACCCCGACGTCACTCATATGATGTA 1902
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 484 GlnArgLeuSerIleAlaArgGlyLeuPheLeuAsnAsnProIleuIleLeuAspGly 503
Qy 1903 GCTACTTCATCGTTAGATTCGATTCGATTACTGAAAGACTATTTCTTGTCGSCATGAAGAGTGTG 1962
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 504 AlaThrSerAlaLeuAspLeuGlnSerGlnSerIleAlaGlnGluAlaLeuAspValLeu 523
Qy 1963 GTCAACACACAGACTTTTATTTATTCATTCGACACACAGATTGTCAACAGTGTGATGCAGAT 2022
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 524 SerIylsAspArgThrIleThrIleValAlaIleHisArgLeuSerThrIleThrHisAlaAsp 543
Qy 2023 GAATACATTTGTCTTGATTCAGAGGGTAAAGTATCCGAACGTGTATCCACCATGCTTTGCT 2082
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 544 LysIleValValIleGluAsnGlyHisIleValAlaGlyThrGlyThrHisIleArgGlnLeuIle 563
Qy 2083 GCTAACCTCATAGTATCTATTCAGAAATGTGGCGCATACACAGAGC 2127
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 564 AlaIyls---GlnGlyAlaIlyrGlnHisLeuIlyrSerIleGlnAsn 577

RESULT 6
US-09-738-626-6020
; Sequence 6020, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUKIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626

```

```

: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 6020
: LENGTH: 621
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-6020

Alignment Scores:
Pred. NO.: 2,356-53 Length: 621
Score: 685.00 Matches: 133
Percent Similarity: 50.23% Conservative: 129
Best Local Similarity: 30.11% Mismatches: 247
Query Match: 16.10% Indels: 72
DB: 9 Gaps: 19

AF133659 (1-2345) x US-09-738-626-6020 (1-621)
OY AAAATCATAAAAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGCGCAATCTACAGACT 414
DB 6 ArgIleuulnysftr---ThrsAlaleutrrProTyrTyleuGlYlleIleVal 24
OY 415 AGAGTGGCACTTTGCGTGGGATTTTGGGTGTGCAAAAGGCCATGAATATTGTGTTCC 474
DB 25 SerIleValIleAla-----AlaleuSerIleuSerPro 36
OY 475 TTCAATGTTAAATATGCTGTAGACAGCCTC--AACCAATGTGGGAAACATGCTGAC 531
DB 37 PheIleuulnAgGluAlaThrAspSerIleValSerAlaValThrGlySer-----Asn 54
OY 532 CTGAGTGATGCA---CCAAATACAGTTGCAACCATGGCAACAGCACTTGATGCTGTA 588
DB 55 ThrAlaAspAlaValThrArgThrIleIlePheIleuAlaIleuAlaIleuPheVal 72
OY 589 GGTGATATCAGACGCTGAGCTGCTTTTTTTAAACGAAGTTGGAATCCAGTA----- 639
DB 73 -----AlaSerPheIleuAsnThrValMetThrAsnIleGlyIleTyr 86
OY 640 TTGGCAAGTAGCCGCAAGATTCATCCGAAGAT--GCCAAATATGCTTTCATCC 696
DB 87 IleGlyAspValMetAlaSerArgMetArgGlnIleuulnIaThrArgTyrTyrAlaIys 106
OY 697 CTTCACACACTGATCTGGGTTTTCACCTGAGACAGACAGCGGAGCTTATCTAAGCT 756
DB 107 LeuIleuAlaIeuProGlnIuIyTyrPheAspAsnGlnValThrGlyThrIleIleAlaIys 126
OY 757 ATTGACAGAGAACACAGGGTATCAGTTTGTCTCTAGAGCTTGGTATTTAATCTTCT 816
DB 127 LeuAspArgSerIleAsnGlyIleThrGlnPheMetClnSerPheSerAsnSnphePhe 146
OY 817 CCCAATC-----AGTTTGAAGTATGCTGTGACAGTGTGTTTGTATTACAAATGC 867
DB 147 PheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 166
OY 868 GGTGCCCAAGTTTGCTTGTGTAACCTTGGAAACACTTGATACATACACAGATTCACAGTT 927
DB 167 Ala-----IleIleuAlaMetLeuPheProIleuTyrMetTyrLeuThr 181
OY 928 GCAATCA-----CGGTGAGACATAGATTGAATGAATGAACAAAGACAGATTAT 981
DB 182 AlaIeuThrSerIyAsaGrrpGln--LysTyrGluGlyGlnIyAsaSnIsgIuIleAsp 200
OY 982 GATCCAGGATATGCTGCTATAGCACTCACTGCTGAATTAGAAACCTTGAGATTTTAAT 1041
DB 201 ValAlaAsnGlyArgPheAlaGluValAlaGlyGlnValIyValValIySerPheVal 220

```

```

Oy 1042 AATGAAGATATGAAGCAGATATGTATGA---TTTTGAAGCGTATGAGACTGCT 1098
Db 221 AAlagIuThArGluLeuAlaAspHeGlyGlyArgTyGlyLeuThrValAlaIleThr 240
Oy 1099 TCAATTGAAAAGT-----ACCTCTACTCTGGCTATGCTGAACTTTGGTCAA 1143
Db 241 ArgProGlnSerGlyTrpTrpHisArgMetAspThrLeu-----Arg 254
Oy 1144 AGTGCATATTTTGGCTTTTAACAGCTATATATGCTATATGCTGCCAGTCAGGAATTG 1203
Db 255 GAlaAlaAlaLeuAsnIleIlePheLeuAlaIleHisLeuLeuIlePheTyArgThrLeu 274
Oy 1204 GCAGGTACCTTACTGTTGGAATCTAGTAATGTGAATGAGCTGCTTTTTCAGCTTTCA 1263
Db 275 HisGlyHisPheThrIleGlyAspMetValMetLeuIleGlnLeuValThrMetAlaGln 294
Oy 1264 TTAACCCCTGAATCTTGTGGAACTGTATATAGAGACTAGACAGCACTCATAGATATG 1323
Db 295 GlnProValTyMetMetSerTyRileValAspSerAlaGlnArgAlaIleAlaGlySer 314
Oy 1324 AACACCTTGTACTCTACTC-----AGGTAGACACCCAAATTAA 1365
Db 315 ArgAspTyRPhGluValMetAlaGlnValGluProThrAlaAsnLysGluLeuVal 334
Oy 1366 GACAAAGTATGCAATCT-----CCCTTCAGATCACA 1398
Db 335 AspAlaThrLeuAlaSerAspThrProArgIleSerValGlyThrProAlaAlaLeuPro 354
Oy 1399 CCACAGACAGCTACCGTCCCTTTGATATATGTGCATTTTGAATACCTTAGAGCCAGAAA 1458
Db 355 AlaGlyLupProAlaMetGlnPheLysAsnValThrPheAlaTyGlnLupLysPro 374
Oy 1459 GTCCCTTAGATATATCCCTTGAAGTCCCTGCAGGAAAGAAAGTGGCCATTGAGAGGT 1518
Db 375 ValIleSerAspValSerIleThrAlaArgHisGlyGlyArgIleAlaLeuValGlyGlu 394
Oy 1519 AGTGGTCAAGGAAAGCAACAATAGTAGAGCTATATTTCCTTATGAGCTCAAAAG 1578
Db 395 SerGlyGlyGlyLysSerThrLeuValAsnLeuLeuGlyLeuTyRlysProAsnSer 414
Oy 1579 GGTACATTTATCTGCTGCTGCAAAATATACAGATGTAGAGCTGGAAGCTTCGGAGG 1638
Db 415 GlySerLeuAlaValCySGlyValAspValLysAspLeuThrSerGlnLupLeuArgAla 434
Oy 1639 GCAGTGGAGGTGATCCTCAGAGTGTGCTCTTCATATATCTTATTATACACCTC 1698
Db 435 SerValGlyValValPheGlnAspAlaSerLeuPheSerGlySerIleAlaGluAsnIle 454
Oy 1699 TTATATGGAACATCAGTCTTCACTCAGAGAGATGTATGAGTGGCAAAATTAGCTGGA 1758
Db 455 AlaTyRGlYArgProGlyAlaThrArgGluGluIleIleGluValAlaIleLysAlaAsn 474
Oy 1759 CTTTCATGATGCATTTCTTGAATGCGACATGATATGACACCCAGTAGGGAGAGAGA 1818
Db 475 AlaHisGluPheIleSerAlaPheProGluGlyTyRGluthrValValGlyLupArgGly 494
Oy 1819 CTCAGAGCTTTGAGAGGAGAAAGCAAGATAGACATTCAGCAAGCCATTGGAAGAGC 1878
Db 495 LeuLysLeuSerGlyGlyGlnLysGlnArgValSerValAlaArgAlaMetLeuLysAsp 514
Oy 1879 CCCCAGTCATCTCTATGATGAGCTACTTCACTGATGATTGATTCATGAGAGACT 1938
Db 515 AlaProLeuLeuValLeuAspGlnAlaThrSerAlaLeuAspThrLysSerGluGlnAla 534
Oy 1939 ATTCTTGGTGCATGAAGATGTGTCAACACAGAACTTCTATTTCATTGTCACACAGA 1998
Db 535 ValGlnAlaGlyLeuGlnLupLeuMetGluAsnArgThrThrLeuMetIleAlaHisArg 554
Oy 1999 TTGTCAACAGTGTGATCAGATGAATCATGTCTTGTCATCAGGATGAAGGCCAGA 2058
Db 555 LeuSerThrIleAlaGlyValAspThrIleValIleThrIleGlnAsnGlyArgValGluGlu 574
Oy 2059 CGTGTATCCACCATGTTGCTTGCTAACCTTCATATCTATTTCAAGAAATGTGGCAT 2118

```

```

Db 575 ValGlySerProThrGlu---LeuAlaValSerGlyGlyIleTyRSerGluLeuArg 593
Oy 2119 ACACAGACGCCGTGTGCAGACACATGATACCCCAAAATGAGCAAGCAAAAGAAAT 2178
Db 594 LeuThrAsnSerThrAla-----GluAlaAspArgGluArg 605
Oy 2179 ATA 2181
Db 606 Leu 606

RESULT 7
US-09-953-259-2
; Sequence 2, Application US/09953259
; Patent No. US20020115159A1
; GENERAL INFORMATION:
; APPLICANT: FARMICK, Mike
; APPLICANT: HUTTMACHER, Klaus
; APPLICANT: PFEFFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE ATR61 PROTEIN
; FILE REFERENCE: 213903USOX
; CURRENT APPLICATION NUMBER: US/09/953,259
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: DE10045579.4
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-953-259-2

Alignment Scores:
Pred. No.: 2,356-53 Length: 621
Score: 685.00 Matches: 193
Percent Similarity: 50.23% Conservative: 129
Best Local Similarity: 30.11% Mismatches: 247
Query Match: 16.10% Indels: 72
DB: Gaps: 19

AF133659 (1-2345) x US-09-953-259-2 (1-621)
Oy 355 AAAATCATAAAGCAATCTTCTTATGTGTGCCCCAAAGACAGGCCAGATCTACGAGCT 414
Db 6 ArgIleLeuLysThr---ThrSerAlaLeuThrProTyRTrpLeuGlyIleIleValVal 24
Oy 415 AGAGTGGCAATTCGCTGCGATTTTGGGTGCTGCAAGGCCATGATATATTGGTCCC 474
Db 25 SerIleValIleAla-----AlaLeuSerLeuLeuSerPro 36
Oy 475 TTTCATGTTAAATATGCTGTAGACAGCTC---AACCAAGATGGGGAAACATGCTGAC 531
Db 37 PheIleLeuArgGluAlaThrAspSerIleValSerAlaValIleGlySer-----Asn 54
Oy 532 CTGAGTGATGCA---CCAAATACAGTTGCAACATGCGCAACAGAGCTTCTGATGGCTAT 588
Db 55 ThrValAspAlaValAlaThrArgThrIleIlePheLeuAlaLeuAlaLeuPheVal----- 72
Oy 589 GGTGTATCAAGAGCTGAGCTGCTTTTATTACGAAGTTGCAAAATGCAAGTA----- 639
Db 73 -----AlaSerPheLeuAsnThrValMetThrAsnIleGlyTyR 86
Oy 640 TTGGCAAGGTAGCCCAAGATTCATCCGAGGATA---GCCAAAAATGCTTCTCCAT 696
Db 87 IleGlyAspValMetAlaSerArgMetArgGlnIleLeuAlaThrArgTyRTrpAlaLys 106
Oy 697 CTTCAACACTGATCTGTGGTTTTCACCTGAGACAGACAGCGGAGCTTATCTTAAGGCT 756
Db 107 LeuLeuAlaLeuProGlnLysTyRPhAspAsnGlnValThrGlyThrIleIleAlaArg 126
Oy 757 ATTGACAGAGGAACAGGGATACGTTTGTCTCTGAGTGCTTTGGTATTTAAATCTTCTT 816

```

```

Db 127 LeuaspargSerIleAsnGlyIleThrGlnPheMetGlnSerPheSerAsnAsnPhePhe 146
Qy 817 CCATC-----ATGTTGAAGTGAAGTGTCTGCTGAGTGGTGTGTTGTTATTAACAATGC 867
Db 147 ProMetIleuIleThnMetValAlaValLeuIleIleSerAlaIlePheTyrTrpProIeu 166
Qy 868 GGTGCCAGTTGGCTTGGTGAACCTTGGAAACCTGCTGATACACACGATTCACAGCTT 927
Db 167 Ala-----IleLeuIleValIleMetLeuPheProIleTyrMetTrpLeuThr 181
Qy 928 GCAGTCACA-----CGGTGAGAGAACTAGATTGAATGAATGAATGAATGAATGAATGAAT 981
Db 182 AlaLeuThrSerLysArgTrpGln--LysTyrGlnGlyGlnLysAsnHisGlnIleAsp 200
Qy 982 GATCAGAGTAATGCTGCTATGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
Db 201 ValAlaAsnGlyArgPheAlaGlnValValGlnValLysValValLysSerPheVal 220
Qy 1042 AATGAAGATATGAAGACACAGATATGATGA--TTTGAAGACGATATGAGCTGCT 1098
Db 221 AlaGlnThrArgGlnLeuAlaAspPheGlyArgTyrGlnLysThrValAlaIleThr 240
Qy 1099 TCATTGAAAGT-----ACCTTACTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
Db 241 ArgProGlnSerGlyTrpTrpHisArgMetAspThrLeu-----Arg 254
Qy 1144 AGTCTATTTTCAGTGTGCTGCTTAAACGCTATTAATGCTGCTGCTGCTGCTGCTGCTGCT 1203
Db 255 GlyAlaIleLeuMetIleIlePheLeuAlaIleHisLeuLeuIlePheTyrArgThrLeu 274
Qy 1204 GCAGGTACCTTACTGTTGAGATCTAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263
Db 275 HisGlnHisPheThrIleGlyAspMetValMetLeuIleGlnLeuValThrMetAlaGln 294
Qy 1264 TTACCCCGAACTTTCTGGGAACTGTATATAGAGACTATACACAGACTCATCATGATATG 1323
Db 295 GlnProValLysMetMetSerTyrIleValAspSerAlaGlnAlaIleGlySer 314
Qy 1324 AACACCTGTTTACTTACTACT-----AAGTACACACCAATTAATA 1365
Db 315 ArgAspTyrPheGlnValMetAlaGlnGlnValGlnProThrAlaAsnLysGlnLeuVal 334
Qy 1366 GACAAAGTATGCACTT-----CCCTTCAGATCACA 1398
Db 335 AspAlaThrLeuAlaSerAspTyrProArgLysSerValGlyThrProAlaIlePro 354
Qy 1399 CCAAGACAGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
Db 355 AlaGlyLysProAlaMetGlnPheLysAsnValThrPheAlaTyrGlnGlnLysPro 374
Qy 1459 GTCCTTAGTGAATATCTTGAAGTCCCTCAGAGAAAGAGGCGCATGTTGAGAGT 1518
Db 375 ValIleSerAspValSerIleThrAlaArgHisGlyGlnAlaGlyIleAlaValGlyGln 394
Qy 1519 AGTGGGTCAAGGAAACACAAATAGTGAAGCTATTATTTGCTTCTGAGCTCAAAAG 1578
Db 395 SerGlyGlyLysSerThrLeuValAsnLeuLeuGlnLysLeuTyrLysProAsnSer 414
Qy 1579 GGTAGCATTTATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638
Db 415 GlySerLeuAlaValCysGlyValAspValLysAspLeuThrSerGlnGlnLeuAla 434
Qy 1639 GCAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1698
Db 435 SerValGlyValValPheGlnAspAlaSerLeuPheSerGlySerIleAlaGlnAsnIle 454
Qy 1699 TTATATGAAACATCACTGCTTCACTGAGAGAGTATGACAGTGAAGTGAAGTGAAGTGAAGT 1758
Db 455 AlaTyrGlyArgProGlyAlaThrArgGlnGlnIleIleGlnValAlaLysValAsn 474
Qy 1759 CTTCATATGCAATTTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1818
Db 475 AlaHisGlnPheIleSerAlaPheProGlnGlyTyrGlnThrValValGlyGlnArgGly 494

Qy 1819 CTCAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878
Db 495 LeuLysLeuSerGlyGlnGlnLysGlnArgValSerValAlaArgAlaMetLeuLysAsp 514
Qy 1879 CCCCAGTACTACTCTATGATGAAGTACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
Db 515 AlaProLeuLeuValIleAspGlnAlaThrSerAlaLeuAspThrLysSerGlnAla 534
Qy 1939 ATTCTTGCTGCAAGAGATGCTGCTCAACAGAGACTTATTTATTCATTCATGACACAGA 1998
Db 535 ValGlnAlaGlyLeuGlnLeuMetGlnAsnArgThrThrLeuMetIleAlaHisArg 554
Qy 1999 TTGTCACAGTGTGCTGATGAGATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2058
Db 555 LeuSerThrIleAlaGlyValAspThrIleValThrIleGlnAsnGlyArgValGlnGln 574
Qy 2059 CGTGTACCCACCATGCTTGTCTGCTTCACTTCACTGATATCTATTCAGAAATGTCAT 2118
Db 575 ValGlySerProThrGln--LeuAlaValSerGlyGlyIleTyrSerGlnLeuLysArg 593
Qy 2119 ACACAGAGCAGCCGTGTCAGAACCATGATACCCCAATGGAGAGAGAGAGAGAGAGAT 2178
Db 594 LeuThrAsnSerThrAla-----GlnAlaAspArgGlnArg 605
Qy 2179 ATA 2181
Db 606 Leu 606

RESULT 8
US-09-815-242-14006
; Sequence 14006, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zysek, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14006
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14006

Alignment Scores:
Pred. No.: 1,4e-50 Length: 582
Score: 654.00 Matches: 187
Percent Similarity: 49.03% Conservative: 115

```

Best Local Similarity: 30.36% Mismatches: 266
 Query Match: 15.37% Indels: 48
 DB: 10 Gaps: 15

AF133659 (1-2345) x US-09-815-242-14006 (1-582)

```

QY 310 CTCACACAGACCCAAAGAGGTTAAAGATGTTACTCGGAATAATCAATAAGCA 369
   |||||
Db 1 MethSAsnAap-----LysAspLeuSerThrTrpGlnThrPheArgArg 15
QY 370 ATGCTTTCTATATGTGGCCCAAGACAGCCAGATCTACAGGCTAGAGTTGCCATTGC 429
   |||||
Db 16 Leu-----TrpProThrIleAlaProPheLysAlaGlyLeuIleValAlaGly 31
QY 430 CTGGGATTTTGGTGGTGGCAAGGCGCATGATTTGGTCCCTTCATGTTAAATAT 489
   |||||
Db 32 IleAlaLeuIleLeuAsnAlaAlaSerAspThrPheMetLeuSerLeuLeuysProLeu 51
QY 490 GCTGTAGACAGCCTCAACACAGTGTGGGAAACATGCTGAACCTGATGTCACCAAT 549
   |||||
Db 52 LeuAspAspGlyPheGlyLys-----ThrAspArg 61
QY 550 ACAGTTGCACCATGCAACACAGACTTGTGATGGCTATGTTCAAGAGCTGAGCT 609
   |||||
Db 62 SerValIleLeuTrpMetProLeuValAlaIleGlyLeuMetIleLeuArgGlyIleThr 81
QY 610 GCTTTTAAAGAA---CTTGAATAGCATTTTGGCAGGTAAGCCAGAAATTCATC 666
   |||||
Db 82 SerTyrIleSerSerTyrCysIleSerTyrPalaSerGlyLysValValMetThrMetArg 101
QY 667 CGAAGAAATAGCCAAATATGCTTCTCCATCTTCAACACCTGATCTGGGTTTCACTG 726
   |||||
Db 102 ArgArgLeu-----PheGlyHisMetMetGlyMetProValAlaPhePheAsp 117
QY 727 ACCAGACAGACGGAGACTTATCTAAGCTATTTACAGAGAAACAAGGGGTATCACTTT 786
   |||||
Db 118 LysGlnSerThrGlyThrLeu-----LeuSerArgIleThrTyrAspSerGlnGln 134
QY 787 GTCCGAGAGCTTGGTATTTAATCTCTCCATCATGATTTGAAGTGAAGCTTGAGCT 846
   |||||
Db 135 ValAlaSerSerSerSerGlyAlaLeuIleThrValAlaArgGlnGlyAlaSerIleIle 154
QY 847 GGT-----GTTTGTATTACAATGCGGTGCCAGTTGCTTGTAAACCTT 894
   |||||
Db 155 GlyLeuPheIleMetMetCHeTyrTyrSerTrpGlnLeuSerIleIleLeuValValLeu 174
QY 895 GGAACACTTGTGATCATCACAGCATTCACAGTTCACAGCGGTGAGAACTAGATT 954
   |||||
Db 175 AlaProIle-----ValSerIleAlaIleArgValValSerLysArgPhe 189
QY 955 AGAATAGAAATGAACAAACAGATATGATGAGGT-----AATGCTGCTATAGAC 1005
   |||||
Db 190 ArgSerIleSerLysAsnMetGlnAsnThrMetGlyGlnValThrThrSerIleGlnGln 209
QY 1006 TCACGTGTAATTAATGAACCTGTAAGTATTTAATTAATGAAGATATGACACAGAGA 1065
   |||||
Db 210 MetLeuLysGlyHisLysGlnValIleuIlePheGlyGlnGlnValGlnThrLysArg 229
QY 1066 TATGATGATTTTGAAGACGATGAGACTGCTTCAATGAAGAAAGTCACTTACTGCT 1125
   |||||
Db 230 PheAspLysValSerAsnLysMetArgLeuGlnGlyMetLysMetValSerLysSer 249
QY 1126 ATGCTGAATTTGGTCAAGGTCTATTTTCACTGTCGTTTAAACAGCTATATGGCTC 1185
   |||||
Db 250 IleSerAspProIleIleGlnLeuIleAlaSerLeuAlaLeuAlaPheValLeuYrAla 269
QY 1186 GCCAGTACGAGGAATTTGTGAGAGTCTTACTGTTGAGTCACTGATTAATGGAATGA 1245
   |||||
Db 270 AlaSerPheProSerValIleMetAspSerLeuThrAlaGlyThrIleThrValAlaPheSer 289
QY 1246 CTGCTTTTCACTTCACTTACCTCCCTGAACCTTCTGGGAACGTATATAGAGACTAGA 1305
   |||||
Db 290 SerMetIleAlaLeuMetArgProLeuLysSerLeuThrAsnValAsnAlaGlnPheGln 309
  
```

```

QY 1306 CAAGCACTATAGATATGAACACCTTGTATTACTTACTCAAGGTAGACACCCAAATTA 1365
   |||||
Db 310 ArgGlyMetAlaAlaCysGlnThrLeuPheAlaIleLeuAspSerGln---GlnGlnLys 328
QY 1366 GACAAAGTATGATGATCTCCCTTCAGATTCACACCCACAGACGTACCCTGCTTGAT 1425
   |||||
Db 329 AspGlu-----GlyLysArgValIleAspArgAlaThrThrLysAspLeuGluPheArg 345
QY 1426 AATGTGATTTTGAATCATATGAGAGGCGCAAGATC-----CTTAGGAAATATCCTTT 1479
   |||||
Db 346 AsnValThrPheThrTyr---ProGlyArgGlnValProAlaLeuAsnIleAsnLeu 364
QY 1480 GAAGTCCCTGCAGAGAAAGAAAGTGGCCATTGTAGAGAGTATGGGTCAAGGAAAGACA 1539
   |||||
Db 365 LysIleProAlaGlyLysThrValAlaLeuValGlyThrSerGlySerGlyLysSerThr 384
QY 1540 ATAGTAGGCTATATTTTCCCTTCTATAGAGCTCAAAAAGGTAGCATTTATCTTGCTGT 1599
   |||||
Db 385 IleAlaSerLeuIleThrArgPheThrAspIleAspGlnGlyHisIleLeuMetAspGly 404
QY 1600 CAAATATACAAAGTGTAGCTGGAAGCCTTGGAGGCGCAGTGGAGGTGCTCTCAG 1659
   |||||
Db 405 HisAspLeuArgGlnTyrThrLeuAlaSerLeuArgAsnGlnValAlaLeuValSerGln 424
QY 1660 GATGCTGCTCTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1719
   |||||
Db 425 AsnValHisLeuPheAsnAspThrValAlaAsnAsnIleAlaTyrAlaArgThrGlnGln 444
QY 1720 ---TCACCTGAGGAAGGTATGACAGTGGCAAAATTAAGTGCATTCATGATGCAATTCTT 1776
   |||||
Db 445 TyrSerArgGlnGlnIleGlnGlnAlaAlaArgMetAlaTyrAlaMetAspPheIleAsn 464
QY 1777 CGAATGCCACATGATATGACCCCAAGTAGGGAAGAGAGACTCAAGCTTCAGGAGGA 1836
   |||||
Db 465 LysMetAspAsnGlyLeuAspThrIleIleGlyGlnGlnGlyValLeuLeuSerGlyGly 484
QY 1837 GAAAGAGAAAGATGACATTTGCAAGACCCATTTTGAAGACCCCCAGTCACTCTAT 1896
   |||||
Db 485 GlnArgGlnAlaGlyIleAlaIleAlaArgAlaLeuLeuArgAspSerProIleLeuIleLeu 504
QY 1897 GATGAGCTACTTCACTGATGATTCATGATCTGACAGACATATTTCTTGTCCTCAAG 1956
   |||||
Db 505 AspGlnAlaThrSerAlaLeuAspThrGlnArgAlaIleGlnAlaIleLeuAsp 524
QY 1957 GATGCTGCAACACAGACCTTATTTTCACTGACACAGATTTGTAACAGTGGTGTAT 2016
   |||||
Db 525 GlnLeuGlnLysAsnArgThrSerLeuValIleAlaHisArgLeuSerThrIleGlnGln 544
QY 2017 GCAGATGAATCATTTGCTTGCATCAGGCTAAGGTAGCCGAGCTGATCCCAACATGCT 2076
   |||||
Db 545 AlaAspGlnIleValAlaValGlnAspGlyIleLeuValGlnArgGlyThrHisSerGln 564
QY 2077 TTGCTGCTAAACCTCATATGATCTATTCAGAAATGTGGCATACAG 2124
   |||||
Db 565 LeuLeuAlaGln---HisGlyValTyrAlaGlnLeuHisLysMetGln 579
  
```

RESULT 9
 US-09-815-242-10115
 ; Sequence 10115, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zykkind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: EITRA.011A

```

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10115
; LENGTH: 582
; TYPE: PR
; ORGANISM: Escherichia coli
US-09-815-242-10115

Alignment Scores:
Pred. No.: 2,61e-50      Length: 582
Score: 651.00           Matches: 190
Percent Similarity: 49.19%      Conservative: 115
Best Local Similarity: 30.65%    Mismatches: 259
Query Match: 15,308             Indels: 56
DB:                               Gaps: 17

AF133659 (1-2345) x US-09-815-242-10115 (1-582)

QY 310 CTCACACAGACCCAAAGAGGTTAAAGATGTGATCTCGCAAAATCATATAAGCA 369
   |||:|||||
Db 1 Methisambsp-----LysaspLeuSerThrTrpGlnThrPheArgArg 15

QY 370 ATGCTTTTATGTGTGGCCCAAGACAGCCAGAT-----CTACGAGCTAGA 417
   |||:|||||
Db 16 Leu-----TrpProThrIleAlaProPheLysAlaGlyLeuIleValAlaGly 31

QY 418 GTTCCATTTCGCTGGAGATTTTGGGTGGTCAAAAGCCATGAATATGTGGTCCCTTC 477
   |||:|||||
Db 32 ValAlaIleuIleLeuAnaAlaIleSerAspThrPheMetLeuSerLeuLeuLysProLeu 51

QY 478 ATGTTTAAATATGCTGTAGACAGCCTCAACCAAGATGTCCGAAACAATGCTGAACCTGACT 537
   |||:|||||
Db 52 Leu-----AspAsp-----GlyPheGly 57

QY 538 GATGCACCAATACAGTTGCAACCATGCAACAGACGATTCGATTGGCTATGTGTATCA 597
   |||:|||||
Db 58 LysThrAspArgSerValIleuValTrpMetProLeuValIleGlyLeuMetIleLeu 77

QY 598 AGAGCTGAGACTGCTTTTAAAGAA---GTTGCAATGACGATATTTCGCAAGGTAGCC 654
   |||:|||||
Db 78 ArgGlyIleThrSerTyrValSerSerTyrCysIleSerTrpValSerGlyLysValVal 97

QY 655 CAGAAATTCATCCGAAGAATAGCCAAAATGCTCTTCTCCATCTTCAACACCTGGATCTG 714
   |||:|||||
Db 98 MetThrMetArgArgArgLeu-----PheGlyHisMetMetGlyMetProVal 113

QY 715 GGTTTACCTGAGCAGACAGAGCGAGCTTTATCTAAGGTATGACAGAGAACAAG 774
   |||:|||||
Db 114 SerPheAspLysGlnSerThrGlyThrLeu-----LeuSerArgIleThrTyr 130

QY 775 GGATACAGTTTGTCTGAGTGTCTTGTGATTATCTTCTCCATCATGTTGAAGTG 834
   |||:|||||
Db 131 AspSerGlnGlnValAlaSerSerSerSerGlyAlaLeuIleThrValValArgGlnGly 150

QY 835 ATGCTTTCAGTGT-----GTTTGTATTACAAAGCGGTGCCAGTTTGCT 882
   |||:|||||
Db 151 AlaSerIleIleGlyLeuPheIleMetMetPheTyrTyrSerTrpGlnLeuSerIleIle 170
```

```

QY 883 TTGGTAACCTTGGAACACTTGGTATACACAGCAATTCACAGTTGACGTACAGCGTGG 942
   |||:|||||
Db 171 LeuIleValLeuAlaProIle-----ValSerIleAlaIleArgValVal 185

QY 943 AGAATCTAGATTTCATGAAATGAACAAAGCATATATGATGACGT-----AAT 993
   |||:|||||
Db 186 SerLysArgPheArgAsnIleSerLysAsnMetGlnAsnThrMetGlyValAlaThrThr 205

QY 994 GCTGTATACATCATCGCTGCAATTTAATGAAACGTGAAGTATTTTAATATGAAGAATAT 1053
   |||:|||||
Db 206 SerAlaGlnIleMetLeuLysGlyHisLysGlnValLeuIlePheGlyGlnGlnVal 225

QY 1054 GAAGACAGAGATATGATGATTTTGAAGCCATGAGACTGCTTCATGAAAGTACC 1113
   |||:|||||
Db 226 GlnThrLysArgPheAspLysValSerAsnArgMetArgLeuGlnIleMetLysMetVal 245

QY 1114 TCTACTGTGCTTATGCTGAACCTTGGTCAAAAGTCTATTTCATGTCGTGTTAAACGCT 1173
   |||:|||||
Db 246 SerAlaSerSerIleSerAspProIleIleGlnLeuIleAlaSerLeuAlaLeuAlaPhe 265

QY 1174 ATATGGTCTGCGCAGTCAAGGAAATTTGGCAGTACCTTACTGTGTGAGATCTAGTA 1233
   |||:|||||
Db 266 ValLeuTyrAlaAlaSerPheProSerValMetAspSerLeuThrAlaGlyThrIleThr 285

QY 1234 ATGTGAATGACCTGCTTTTTCAGCTTTTCATTCACCTTGAACCTTTCGGAACGTATAT 1293
   |||:|||||
Db 286 ValValPheSerSerMetIleAlaLeuMetArgProLeuLysSerLeuThrAsnValAsn 305

QY 1294 AGAGACATGACAGAACACTCATGATATGAAACACTTGTTTACTTACTCAAGTAGAC 1353
   |||:|||||
Db 306 AlaGlnPheGlnArgGlyMetAlaIleCysGlnThrLeuPheThrIleLeuAspSerGln 325

QY 1354 ACCCAATTTAAACAAAGATGATGCATCTCCCTTCAGATCACACACAGACGATACC 1413
   |||:|||||
Db 326 --GlnGlnLysAspGln-----GlyLysArgValIleGlnArgAlaThrGlyAsp 341

QY 1414 GTGGCTTGTATATGTGACATTTTGAATACATGAGGCGCAAGATC-----CTTAGT 1467
   |||:|||||
Db 342 ValGlnPheArgAsnValThrPheThrTyr---ProGlyArgAspValProAlaLeuArg 380

QY 1468 GGAATATCTTTGAAGTCCCTGCGAGAAAGAGTCCATTTGAGAGTATGGGCTCA 1527
   |||:|||||
Db 361 AsnIleAsnLeuLysIleProAlaGlyLysThrValAlaLeuValGlyArgSerGlySer 380

QY 1528 GGGAAAGCAATATAGTAGGCTTATTTGCTTTTATAGCCTCAAAAAGGTAGCATT 1587
   |||:|||||
Db 381 GlyLysSerThrIleAlaSerLeuIleThrArgPheTyrAspIleAspGlnGlyGlnIle 400

QY 1588 TATCTTGCTGTCGCAAAATATATACAAAGATGTGAGCTGGAAGCCTTGGAGGCGCAGTGGGA 1647
   |||:|||||
Db 401 LeuMetAspGlyHisAspLeuArgGlyTyrThrLeuAlaSerLeuArgAsnGlnValAla 420

QY 1648 GTGGTACTCAGAGATCTGCTCTTCATTAATATCTATTATTAACAACCTTATATGGA 1707
   |||:|||||
Db 421 LeuValSerGlnAsnValHisLeuPheAsnAspThrValAlaAsnMetIleAlaTyrAla 440

QY 1708 AACATCAGTGT---TCACCTGAGAGAGTATGACGTGCGCAAAATTAGCTGACCTTCAT 1764
   |||:|||||
Db 441 ArgThrGlnGlnTyrSerArgGlnGlnIleGlnGlnAlaAlaArgMetAlaTyrAlaMet 460

QY 1765 GATGCAATCTTCTGATGCAACATGATGATGACACCCAGATGAGGAGAGAGACGATCAAG 1824
   |||:|||||
Db 461 AspPheIleAsnLysMetAspAsnGlyLeuAspThrValIleGlyLysGlnGlyValLeu 480

QY 1825 CTTTCAGAGAGAGAAAGCAAGATGCAATTTGCAAGGCCATTTTGAAGACCCGCCA 1884
   |||:|||||
Db 481 LeuSerGlyGlyGlnArgGlnArgIleAlaIleAlaArgAlaLeuLeuArgAspSerPro 500

QY 1885 GTCATATCTATGATGACGACTTATCTGATTAGATTACATTAAGCAAGACTATTTCTT 1944
   |||:|||||
Db 501 IleLeuIleLeuAspGlnAlaThrSerAlaLeuAspThrGlnSerGlnArgAlaIleGln 520
```


QY 1801 CAAGTAGGGGAAACGAGCTCAAGCTTTCAGAGAGAGAAAAAGCAAGAGTAGCAATTGCA 1860
:::|||||
Db 478 ValIleGIyGluAenGIyThrSerLeuSerGIyGIyGlnArgLeuAlaIleIa 497
QY 1861 AGAGCCATTGGAAGAGCCCCAGCTACTACTATGATGTAAGAGCTTCACTTCGTTGAT 1920
|||||
Db 498 ArgAlaLeuLeuArgAsnSerProvalLeuLeuAspGluAlaThrSerAlaLeuAsp 517
QY 1921 TCGATTACTGAAGAGACTATCTTGTCCTCAAGATGTCATCAACAGAACTTCT 1980
:::|||||
Db 518 ThrGluSerGluArgAlaIleGlnSerAlaLeuGluLeuLysValAspArgThrVal 537
QY 1981 ATTTTCATGCACACAGATTGTCAACAGTGTGATGCAATGAAATCATTTGTTGAT 2040
:::|||||
Db 538 ValValIleAlaHisArgLeuSerThrIleGluAsnAlaAspGluIleLeuValIleAsp 557
QY 2041 CAGGTAAGTAGACCCGAACTGTGTACCCAGTGTGTTGCTTGTCAACCTCAATGATC 2100
|||||
Db 558 HisGIyGluIleArgGluArgGlyAsnHisLysThrLeuLeu---GluGlnAenGIyAla 576
QY 2101 TATTCAGAAATGTGCATATACAG 2124
||| :|||
Db 577 TyrIyGlnLeuHisSerMetGln 584

RESULT 11
US-09-815-242-10960 ; Sequence 10960, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10960
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-10960

Alignment Scores:
Pred. No.: 1,34e-48
Score: 632.00
Percent Similarity: 49.01%
Best Local Similarity: 30.76%
Query Match: 14.86%
DB: 10
Length: 587
Matches: 187
Conservative: 111
Mismatch: 264
Indels: 46
Gaps: 17

AF133659 (1-2345) x US-09-815-242-10960 (1-587)
QY 340 GATGTTGATCTCGGAAATCATATAAGCAATGCTTTCTTATGTGTGCCCCAAGACAG 399
||| :|||
Db 10 AspPheSerThrLeuGlnThrPheLysArgLeu-----TrpPheMetIleLys 25
QY 400 CCAAGTATACAGCTAGAGTTGCCATTTCCGCTGGATTTTGGTGGTGCMAAGCCATG 459
||| :|||
Db 26 Pro---PheLysAlaGlyLeuIleValSer-----GlyValAlaLeuValPhe 40
QY 460 AATATGTGTTCC-----TTGATGTTTAATATGCTGTAGACAGCTTCAAC 507
||| :|||
Db 41 AsnAlaLeuAlaAspSerGlyLeuIleTyrLeuLeuLysProLeuLeuAspAsp----- 58
QY 508 CAGATGCGGAAACATGCTGAACTGAGTATGACCAACCAATACAGTTGCAACATGCA 567
||| :|||
Db 59 -----GlyPheGlyLysAlaAsnHisSerPheLeuLysMetMet 71
QY 568 ACAGACTTGTATGCTATGCTATGCTATGCAAGAGCTGAGCTGTTTAAAGAAATT 627
||| :|||
Db 72 AlaPheValValAlaGlyMetIleIleLeuArgGlyIleThrAsnPheIleSerAsnTyr 91
QY 628 CGAAATGCA---GATTTGGCAAGTATGCCAGAAATTCATCCGAAATATAGCCAAAT 684
||| :|||
Db 92 CysLeuAlaTrpValSerGlyLysValAlaMetThrMetArgArgArgLeu----- 108
QY 685 GTCCTTCTCCATCTTCAACCACTGATCTGGTTCCTTTCACCTGAGCAGACGGAGCT 744
||| :|||
Db 109 ---PheLysHisLeuMetPheMetProvalSerPhePheAspGlnAsnSerThrGlyArg 127
QY 745 TTAATCAGGCTATGTGACAGAGACAGAGAGGATATGATTTGCTGAGCTTTGGTA 804
||| :|||
Db 128 Leu-----LeuSerArgIleThrTyrAspSerGlnMetIleAlaSerSerSer 144
QY 805 TTTAATCTTCTCCATCATGTTGAA-----GTATGCTTGTACAGTGTGTTTG 855
||| :|||
Db 145 GlySerLeuLeuThrIleValArgGluGlyAlaTyrIleLeuSerLeuPheAlaValMet 164
QY 856 TATTACAAATGCGGTCGCCAGTTGCTTGGTATACCTTGAACCTGTTGATACATAC 915
||| :|||
Db 165 PheTyrThr-----SerTrpGluLeuThrIleValLeuPheIleIleGlyProIleIle 182
QY 916 GCATTCACAGTTGCAGTACACCGGTGAGAGAACTAGATTAGAAATAGAACAAAGCA 975
||| :|||
Db 183 AlaValLeuLeuArgLeuValSerLysIlePheArg---ArgLeuSerLysAsnLeuGln 201
QY 976 GATTAAT-----GATGAGGTAATGCTGTATAGACTCACTGCTGAATTAAGAACTGTG 1029
||| :|||
Db 202 AspSerMetGlyGluLeuThrSerAlaThrGluGlnMetLeuLysGlyHisValAla 221
QY 1030 AACTATTATTAATGAAGATATGAAGCAGAGATATGATGATTTTGAAGACCTAT 1089
||| :|||
Db 222 LeuSerPheGlyGlnHisValGluGluValHisPheAsnHisValSerAsnAspMet 241
QY 1090 GAGACTGCTTGAAGAAAGTACTACTGCTGCTATGCTGATGCTGCTGCTGCTGCT 1149
||| :|||
Db 242 ArgArgLysSerMetLysMetValThrAlaAsnSerIleSerAspProvalValGlnVal 261
QY 1150 ATTTTCAAGTGTGGTTTAAACAGTATATGCTGCTGCCAGTACGAGGAATTTGGCA--- 1206
||| :|||
Db 262 IleAlaSerIleValLeuAlaThrValLeuTyrLeuAlaThrThrProLeuIleAlaGlu 281
QY 1207 GGTACCTTACTGTGAGATCTAGTAATGAGTAATGAGACTGCTTTTACGCTTCAATTA 1266
||| :|||
Db 282 AspAsnLeuSerAlaGlySerPheThrValValPheSerSerMetLeuAlaMetMetArg 301
QY 1267 CCCCTGAACCTTGTGGGAACTGTATATAGAGACTAGACCAAGCTCATATATGAAC 1326
||| :|||
Db 302 ProLeuLysSerLeuThrLysValAlaValAlaGlnPheGlnSerGlyMetValAlaCysGln 321
QY 1327 ACCTGTTTACTTACTCAAGGTAGACACCCAAATTTAAAGCAAAAGTATGATGCTATCCC 1386
||| :|||
Db 322 ThrLeuPheAlaIleLeuAspLeuGluProGluLysAspAsp-----GlyAla 337

```

Oy 1387 CTTGATCAGACACAGACAGACAGCTACCGTGGCTTTGATATGTCATTTGATACAT 1446
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 338 TTYLYSALAGLPRALALAYSELGLULLEUGLUPHELYASNLVALSERPHEALAYRGLN 357
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1447 ---GAGGGCCAGAAAGTCTTAGTGAATATCTTTGAAGTCCGACAGAAAGAAAGTG 1503
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 358 GLYLSAEPGLULLEUALALEUENASNILLESERPHESERVALPROLAGLYSTRVAL 377
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1504 GCCATTGATGAGGTAGTGGGTGAGGAAAGCAATATGATGAGGCTTATTTGGCTTC 1563
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 378 ALALBUVALGLYARSETGLYSERGLYSERTRILALASNLVALTHIRARGPHE 397
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1564 TATGAGCTTCAAAAGGTAGACATTTATCTTGGTCAAAATATACAGATGTGAGCTTC 1623
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 398 TYRASPILLEGULGNGLYGLULLEULEUENASPGLYVALASNILLEGILNASPYRARGLEU 417
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1624 GAAAGCTTCGGAGGGAGTGGAGTGTACTCAAGATGTCTGCTTCATATACT 1683
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 418 SERASNLVALRGLULASNCYSALVALVALSERGLNGLNVALHILEUPHEASNPTR 437
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1684 ATTATTCACACCTCTTATAT---GGAAACATCAGTCTTCACCTGAGAACTGATGCA 1740
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 438 ILEALAASNLILALATYRVALAGLNASPLYSTRSERARGLULGLULLEILEALA 457
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1741 GTGGCAAAATTGACTGACTCATGATGCAATTTCCGAATGCCACATGATATGACACC 1800
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 458 ALAALALYSALALATYRVALALEUGLUPHEILLEGULYSLEUPROGLNVALPHEASNPTR 477
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1801 CAATGAGGGAAAGAGACTCAAGCTTTGAGAGAGAGAAAGCAAGATGCAATTGCA 1860
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 478 VALILEGLYULASNLGLYSTRSERLEUSERGLYGLYGLNARGLULLEILEALA 497
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1861 AGAGCATTTTGAAGAGACCCCGACATCTATCTATGATGAGCTTATCTTCACTGATGAT 1920
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 498 ACGAALALEULNARGASNSERPROVALLEULILEUASPGILALATHSERXALAEUSP 517
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1921 TCGATTATCGAAGAGACTTCTTGGTCCATGAAGTGTGTCACACAGCACTTCT 1980
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 518 THRGULSERGLULARGALALILEGNLSERLALALEUGLULLEULYALYASPARGTHRVAL 537
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 1981 ATTTTCATTGACACAGATTTGCAACAGTGTGTGATGACAGATGAATCATTTGCTGGAT 2040
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 538 VALVALILEALAHISARGLEUSERTRILLEGULASNLALASPGILILEUVALILEASP 557
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 2041 CAGGATGAGTACCGGAGGTGTACCCACCATGTTGCTTCAACCTCATAGTATAC 2100
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 558 HISGLYGLULLEARGLULARGLYASNHLSTRHLEULEU---GLULNASNLGLYALA 576
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 2101 TATTCGAAATGTGGCATACACAG 2124
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 577 TYRLYSGILNEUHLISERWETGLN 584
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 1410
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 5202
/ LENGTH: 603
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5202

Alignment Scores:
Pred. No.: 2,47e-47 Length: 603
Score: 618.00 Matches: 185
Percent Similarity: 49.02% Conservative: 116
Best Local Similarity: 30.13% Mismatches: 237
Query Match: 14.53% Indels: 76
DB: 10 Gaps: 18

AF133659 (1-2345) x US-09-815-242-5202 (1-603)
Oy 355 AAAATCATMAAAGCAAGCTTTCTTATGTGTGGCCAAAGACAGGCGAGATCTACGAGCT 414
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 14 LYSILETYRPHARGLEULEUGLYTYRVAL-----LYSPROTYLILEGLYMET 29
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 415 AGAGTGGCATTTGCGCTGGGATTTTGGGTGCTGCCAAAGCCATGATATTTGTGCTCCC 474
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 30 PHELEUSERILEVALIGLYPHEULILEPHEALASERTRHGINPROMETLEULAGLY 49
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 475 TTCATGTTAAATATGTGTGATGACAGGCTTCACAGATGTGGGAAACATGCTGAC--- 531
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 50 ILEUEN---LYSTRPHEVALASPGILYLEUSERNSPROASNPALALALEUPHEPROASN 68
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 532 -----CTGAGTATGACACCAATACAGTTGCAACCATGACGCAAGCAGATT 576
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 69 VALGINTPRPTPLLEUARGASPLEUHLISERVALTYRVALVALPROLEULEULILE 88
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 577 CTGATT-----GGCTATGCT-----GATCAAGAGCT 603
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 89 LEULILEALALATRGINGLYLEUGLYSERPHELEUGLYASNPHENPHELEUALALEVAL 108
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 604 GGAGCTGCTTTTAAACGAGTTCGAAATGCAGATTTTGGCAGATGAGCCAGAAATTCA 663
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 109 SERLEUGLYLEUVALHISASPLEUARGVALALEUPHEASNLYSLEULEU----- 125
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 664 ATCCGAAGATAGCCAAATAATGTCTTTCTCCATCTTCAACAACCTGAGT----- 711
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 126 -----VALLEUPROASNARGTYRPHENSPTRHHSERSEGLYHISLEULELSEER 142
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 712 ---CTGGTTTTCACCTGAGACAGACAGCGGAGCTTTATCTAAAGCTATTGACAGAGA 768
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 143 ARGILETHRPHENASNVALTHMETVALTHRGLYALALATHRPARALALILEYVALVAL 162
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 769 ACAAGG---GGTATCAGTTTGTCTGAGTCTTGTGATTTAATCTTCCCATCATAG 825
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 163 ILEARGGLYGLYLEUTH-----VALVALPHELEUPHELEUTYRLEULEU 177
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 826 TTGAGAGATGCTTGCAGTGTGTTTGTATTAACAATCGGTGCCACAGTTGCTTGG 885
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 178 TRP-----METASNPTRYS----- 182
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 886 GTAACCTTGAACACTTGTGATACACAGACATTCACAGTTGAGTACACAGGTTGAGAGA 945
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 183 LEUTHRLEUVALMETLEUALALEUPROVALILEALVALMETVALTHRTHRALASER 202
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```



```

Qy 946 ACTAGATTAGATGAATGAACA-----GCAGATATGATGCAGTAATGCT 996
Db 203 ArglysphearglysglnserlyserileglnvalalmetglAasprvalThrlval 222
Qy 997 GCTATAGCTACCTGCTGATTAATGAACCTGGAAGTATTTATTAATGAAGATATGAA 1056
Db 223 AAserrglthrtrileglnlyTyrrargvalValargserphegylglyalalagtrglu 242
Qy 1057 GCACAGATATGATGATTTTGTGAAGCATGATGATGCTGCTTCAATGAAGTACCTCT 1116
Db 243 Glulysargpheleuaspalaserlserthrlysplysglnleuamethrlyys 262
Qy 1117 ACTGTGCTATGCTGAATTTGTGCAAGTCTTATTTTCACTGCTGTTTAAACAGTATA 1176
Db 263 ThrghlalaValTyrrhPromeleuGlneulvalleTyrralalalmetalalileu 282
Qy 1177 ATGTGCTGCGCAGTACGAGGAATTTGTCAGTACCTTACCTTTGGAGATCTGTAATG 1236
Db 283 MetpheleuValleuTrp-----LeuargglyAspalaSerlaaglyAspleuValala 300
Qy 1237 GTGAATGAGCTTTTTCAGCTTTTCACTTATTAACCTTGAACCTTCTGGAATGTATATGA 1296
Db 301 TyrrlethrAlaalaaglyleuLeuProlysprollearglnleuSerGluValSer 320
Qy 1297 GAGACTAGACAGACCTCATGATATGAACCTGTGTTACTCTACCTCAAGTACAGACACC 1356
Db 321 ThrValGlnarglyValalaglyAlaGlnserlelphelglnleuAsprgluAlala 340
Qy 1357 CAA-----ATTAAACAAAGATGATGAGCATCTCCCTTCAGATCACCA 1401
Db 341 GlnGluuarglnglyThrrValGlnlysergluArgylaserGlnleuValarg--- 359
Qy 1402 CAGACAGTACCGGCTTGATGATGATGATTAATGATTAATGATGATGAGGCGCAGAAA--- 1458
Db 360 -----AsnleuSerPhearglyTyrrProglyThrasprylcIn 371
Qy 1459 GTCCTTAGTGAATATCTTTGAAAGTCCCTGACGAGAAAGAGTGCATTTGAGAGT 1518
Db 372 ValLeuaspapllieserPheilealagluProglyGlnmetleleuValalGlyArg 391
Qy 1519 AGTGGGTAGGGAAGACACATAGTGAAGCTATTAATTTGCTTCTGATGAGCTCAAAAG 1578
Db 392 SerGlyserglylyserThrleuAlaasnleuValProargPheTyrrGlnlelAsnAsp 411
Qy 1579 GGTAGCAATTATCTTGCTGCTGCAAAATATACAGATGAGCTGGAAGGCTCGAGG 1638
Db 412 GlnlyslleleuLeuaspGlyValalGlnuAsprTyrrargLeuargAsnleuArgArg 431
Qy 1639 GCAGTGGAGTGTACCTCAGAGATGCTGCTCTTCCATTAATTAATTAATTAACACTC 1698
Db 432 HistlealeuValThnglnlnValThrlleuPheAsnAspSerValAlaAsnle 451
Qy 1699 TTATATGGAACATCAGTCTTCACT--GAGAAAGTATGAGTGCAGGCAAAATTAAGT 1755
Db 452 AlatyrrGlyAspLeuAlaaglyAlaProargGlnGlnleGlnuArgAlaAlaAla 471
Qy 1756 GGACTCATGATGCAATCTTGAATGCCACATGATGATGACCCCAAGTGGGGAAGCA 1815
Db 472 AsnAlalysGlnuPheleAspAsnleuProglnglyPheAspThrGluValGlyAsn 491
Qy 1816 GGAATCAAGCTTTTGAAGAGAGAAAGCAAGAGTATGCAAGTGCAGGCAATTTGGAAG 1875
Db 492 GlnValleuSerGlyGlnarglnArgleuAlalelAlaGlnAlaLeuLeuyls 511
Qy 1876 GACCCCGAGTCACTATGATGAGCTATCTTCATCGTTAGATTGATTAAGTAAAG 1935
Db 512 AspAlaProleuLeuileleuAspGlnAlaThrserAlaLeuAspThrGlnSerGluArg 531
Qy 1936 ACTATTTGTTGTCAGTGAAGATGGTGTCAACAGAAATTTCTATTTTTCATTTGACAC 1995
Db 532 HistleGlnAlaAlaLeuAspGlnuValMetlyserGlnArgThrThrleuValleleAlaHis 551

```

```

Qy 1996 AGATTGCAACAGTGTGATGCAATGAATCATTTCTTGATCAGGTAAGTACCC 2055
Db 552 ArgleuSerThrtrileglnlysalAspleuileleuValmetAsprglnGlnleVal 571
Qy 2056 GAACGTGTACCCACCATGTTTGCTTGCCT--AACCTCAT 2094
Db 572 GlnarglyserHisAlaGlnleuLeuAlaGlnAsnGlyHis 585

RESULT 13
US-09-749-340-6
; Sequence 6, Application US/09749340
; Patent No. US20020037920A1
; GENERAL INFORMATION:
; APPLICANT: Prichard, Roger K.
; APPLICANT: Xu, Ming
; APPLICANT: Ribeiro, Ana Paula
; APPLICANT: Blackhall, William J.
; APPLICANT: Beech, Robin N.
; APPLICANT: Molento, Marcelo
; APPLICANT: Liu, Hao Yuan
; TITLE OF INVENTION: Methode for Detecting and Reversing
; TITLE OF INVENTION: Resistance to Macrocyclic Lactone Compounds
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,340
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,676
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Renda, Barbara L.
; REGISTRATION NUMBER: 27,626
; REFERENCE/DOCKET NUMBER: 33,333-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-683-2153
; TELEFAX: 973-683-4109
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-749-340-6

Alignment Scores:
Pred. No.: 5,65e-47 Length: 1275
Score: 615.50 Matches: 196
Percent Similarity: 48.01% Conservative: 130
Best Local Similarity: 28.87% Mismatches: 267
Query Match: 14,47% Indels: 86
DB: 10 Gaps: 21

AF133659 (1-2345) x US-09-749-340-6 (1-1275)
Qy 228 GTTCTTAGATGCTGCAAGGCTCTCCAGGTATGCGCATGATGAAGAAGCATGTTG 287
Db 631 ValLeuarglnglnlysglnGlnGlnuValleuAspSerAspAlaGlnSerAspValVal 650
Qy 288 GCATGTGATGACGAGAGGAGTCCACACAGACCCAAAGAGGGTTAAAGATGTTGA 347

```

```

Db      651 -----SerProAsp1leAlaLeuProHisLeuSerSerLeuArgSerArgLys-Glucose 668
Oy      348 TACTGGGAAAATCATATAAAGCAATGCTTTCTAT---GCTGGCCCAAGAACAGCCGCA 404
Db      668 rThrArgSerAlaIleSerAlaValProSerValArgSerMetGlnIleGluMetGluAs 688
Oy      405 TCTAGAGCTAGAGTTGCC----- 423
Db      688 pleuArgAlaLysProThrProMetSerLysIlePheThrPheAsnArgAspLysTrpG1 708
Oy      424 ----ATTTCGCTAGATTTTGGGTGGTGCAGAACGCCATGAATATTGGTTCCTTCAT 479
Db      708 yThrPheIleLeuGlyLeuIleAlaCysIleIleThrGlyThrValThrProThrPheAl 728
Oy      480 GTTTAAATATGCT-----GTAGACAGCCCTCAACCCAGATGCGGAAA 521
Db      728 aValLeuValArgAlaGlnIleIleGlnValLysSerGluProValAspGlnMetLysGlyH1 748
Oy      522 CATGCTGAACCTGATGATGACACCAAAATACAGTTGCACACATGCGCAACAGCAGTTCTGAT 581
Db      748 eValLeuPheTrpCysGlyAla-----Phe11 757
Oy      582 TGGCTATGCTGATACAGAGCTGAGCTGCTTTTAAAGAAATGCGAAATGCACTATT 641
Db      757 eValIleGlyLeuValHisAlaPheAlaPhePheSerAlaIle-----Cyste 774
Oy      642 TGGCAAGTAGCCCAAGATTCATCAATCCGAAGATAGCCAAATGCTTTCTTCATCTTCA 701
Db      774 uGlyArgCysGlyGluAlaLeuThrLysLysLeuArgPheGluAlaPheLysAsnLeu 794
Oy      702 CAACCTGATCTGGGTTTTCAC-----CTGACAGACAGACGAGGAGCTTTATCTAAAGC 755
Db      794 uArgGlnAsnValGlyPheThrAspAspIleArgHisGlyThrGlyLysLeuCysThrArg 814
Oy      756 TATTGACAGAGGAAACAGGGGATATCATAGTTTGTCTGAGTGGTTGGTATTTAATCTTCT 815
Db      814 gPheAlaThrAspAlaProAsnValArgTyr-----ValPheThrArgLe 829
Oy      816 TCCC---ATCATGTTTGAAGTATGCTGTCTCACTGCT-----GTTTGTATACAAAG 866
Db      829 uProGlyValLeuSerSerValValThrIleIleGlyAlaLeuValIleGlyPheIlePh 849
Oy      867 CGGTGCCAGTTTGTCTGTGAACCTTGA-----ACACTGTGATACATAC 914
Db      849 eGlyTrpGlnLeuAlaLeuIleLeuMetValMetValProLeuIleIleGlySer----- 867
Oy      915 ACCATTACAGTTGCACACACGCTGAGAACTAGATTAGAAATGAATGAACAAGC 974
Db      868 -----GlyTyrPheGlnuMetArgMetGlnPheGlyLysLysMetLArgAspThr 883
Oy      975 AGAT-----AATGATGCGATATATCTCTATAGACTACAGCTGCAATATGAACCT 1028
Db      883 rGluLeuLeuGlnGluAlaGlyLysValAlaSerGlnAlaValGluAsnIleArgThrVal 903
Oy      1029 GAAGTATTTTAAATGAAGATATGAACACAGATATGATGATTTTGAAGACGTA 1088
Db      903 lHisAlaLeuAsnArgGlnGluGlnPheHisPheMetTyrCysGluTyrLeuLysGluPr 923
Oy      1089 TGAGACTGCTTGAAGAAAAGTACTCTACTG---GCTATGCTGAACCTTTGGTCAAG 1145
Db      923 cTyrArgGluAsnLeuCysGlnAlaHisThrTyrGlyValPheAlaPheSerGlnSe 943
Oy      1146 TGCTATTTTCACTGCTGTTTAAAGCTTAAATGATGCTGCCAGCTGAGGGAATGTGAC 1205
Db      943 rLeuLeuPhePheMetTyrAlaValAlaPheTrpLys----- 955
Oy      1206 AGGTACCTTACTGTTGGAGATCTAGTATAGTGGAATGAGTGTTCCTTTCAGCTTCATT 1265
Db      956 -GlyAlaIlePheValAspAsnHisSerMetGlnProIleAspValTyrArgValPhePh 975
Oy      1266 ACCCTGAACCTTTCTGGGAATGATATATAGAGACTAGACAAGCACTCATATATAG-- 1323

```

```

Db      975 eAlaPheMetPheCysGlyGlnMetValGlyAsnIleSerSerPheIleProAspValVa 995
Oy      1324 -----AACACTTGTATTACTCTACTACAGGTAGACACCAATTA 1364
Db      995 lLysAlaArgLeuAlaAlaSerLeuLeuPheTyrLeuIleGlnHisProSerGluIleAs 1015
Oy      1365 AGACAAAGTATGATGCTATCTCCCTTCAGATCACACACAG---ACAGCTACCGTGGCTT 1421
Db      1015 pAsn-----LeuSerGluAspGlyValThrLysLysIleSerGlyHisIleSerPh 1032
Oy      1422 TGATTAATGTCATTTTGAATTCATTGAGGCCAG-----AAAGCTCTTACTGGAATATC 1475
Db      1032 eArgAsnValLysrPheAsnTyrProThrArgArgGlnIleArgValLeuArgGlyLeuAs 1052
Oy      1476 CTTGAAGTCCCTCAGAAAGAAAGATGAGCATTTGAGAGTGTGGTGCAGGAAAG 1535
Db      1052 nLeuGlnIleAsnProGlyThrThrValAlaLeuValGlyGlnSerGlyCysGlyLysSe 1072
Oy      1536 CACAATAGTGAAGCTATTATTGCTTCTATAGCTTCAAAGGGTATGACATTATCTTGC 1595
Db      1072 rThrValMetAlaLeuLeuGlnuArgPheTyrAsnGlnAsnLysGlyValIleThrValAs 1092
Oy      1596 TGGTCAAAATATATACAAATGATGAGCCCTGGAAGCCCTCGAGGCGCAGTGGAGTATCC 1655
Db      1092 pGlyGlnAsnIleArgAsnMetAsnIleArgAsnLeuArgGlnuValCysIleValSe 1112
Oy      1656 TCAGATGCTGCTCTCTCCATTAATCTATTATTAACCTCTTATATGGAACATCAG 1715
Db      1112 rGlnGluProThrLeuPheAspCysThrIleMetGlnAsnIleCysTyrGlyLysuAspAs 1132
Oy      1716 TGCTTCACT-----GAGAAAGTATGACGTGCGCAAAATTAAGCTGATCATGATGC 1769
Db      1132 pProLysProSerTyrGlnGlnuValAlaAlaAlaIleLysMetAlaAsnIleHisAsnPh 1152
Oy      1770 AATTCTTGAATGCGCACATGATATGACACCAAGTACGGAAGCAAGCACTCAAGCTTC 1829
Db      1152 eValLeuGlyLeuProGlnGlnuLysrPheThrArgValGlyGlnLysGlyThrGlnLeuSe 1172
Oy      1830 AGGAGGAGAAAGCAAGAGTAGACATTTGCAAGAGCCATTTTGAAGACCCCGCAT 1889
Db      1172 rGlyGlyGlnLysGlnuArgIleAlaIleAlaArgAlaLeuIleLysAspProPheIle 1192
Oy      1890 ACTCTATGATGAAGCTACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1949
Db      1192 uLeuLeuAspGlnuAlaThrSerAlaLeuAspThrGluSerGlnuLysIleValGlnAspAl 1212
Oy      1950 CATGAAGATGTGCTCAACACACAGACTTCTATTTTCAATGTCACACAGATGTCAAAGT 2009
Db      1212 aLeuGlnuValAlaArgGlnuLysrPheThrCysLeuValIleAlaHisArgLeuSerThr11 1232
Oy      2010 GGTGATGACAGATGAATCATTTGCTTGGATCAGGATAGGTAAGGTAAGGTAAGGTAAGGTA 2069
Db      1232 eGlnHisSerSerValIleValIleMetIleGlnGlnuLysValAlaThrAspAlGlyThrHi 1252
Oy      2070 CCATGCTTGTCTGCTCAACCTCATATCTATTCAGAAATGTCACATACAG 2124
Db      1252 eGlnHisLeuLeuMetLys---AsnAspLeuTyrLysArgLeuCysGlnuThrGln 1269

```

```

RESULT 14
US-09-738-626-4751
Sequence 4751, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAKO
APPLICANT: SENOH, AKIHRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

```

```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; PRIORITY FILING DATE: 2000-12-18
; PRIORITY FILING DATE: 2000-12-18
; PRIORITY FILING DATE: 1999-12-16
; PRIORITY FILING DATE: 2000-04-07
; PRIORITY FILING DATE: 2000-04-07
; PRIORITY FILING DATE: 2000-08-03
; PRIORITY FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4751
; LENGTH: 1247
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4751

Alignment Scores:
Pred. No.: 5,33e-45 Length: 1247
Score: 593.50 Matches: 167
Percent Similarity: 48.43% Conservative: 111
Best Local Similarity: 29.09% Mismatches: 265
Query Match: 13.95% Indels: 31
Gaps: 11

AF133659 (1-2345) x US-09-738-626-4751 (1-1247)

QY 418 GTTGCCATTTGGCTGGGATTTTGGTGCGCAAGGCATGATATTGTCCTTC 477
DB 695 ValAlaValAlIleAlaLeuLeuValGlyVal--AlaAlaAspLeuAlaPheProThr 713
QY 478 ATGTTAAATATGCTGTAGACAGCCTCAACGAGATGTCGGAACATGCTGACCTGAT 537
DB 714 LeuMetArgAlaAlaIleLeuAspAsnGlyValGln-----Ala 725
QY 538 GATGCACCAAAATACAGTTGACCAACGAGCAACAGCAGTTGATGCTGATGCTATCA 597
DB 726 GlnSerThrSerThrLeuTrpTrpIleAlaIleAlaGlySerValValLeuLeuSer 745
QY 598 AGACCTGAGCGCTGTTTTTAAAGAACTGGAATGACGATATTGGGAAGTACCAG 657
DB 746 TrpAlaAlaAlaAla-----IleAsnThrIleIleThrAlaArgThrGlyGlu 761
QY 658 AATCAATCCGAGAGATAGCAAAATGCTTTCCTCATCTTCAACCTGATCTGGGT 717
DB 762 ArgLeuLeuArgGlyLeuArgLeuArgSerPheValHisLeuLeuArgLeuSerMetSer 781
QY 718 TTTCACCTGACGACAGACGCGAGCTTTATCTAAGCTATGACAGAGAAAGGGGT 777
DB 782 TyrPheGluArgThrMetSerGly-----ArgIleMetThrArgMetThrThrAsp 798
QY 778 ATCAGTTTTCCTGAGTCTTTGGTATTTATCTTCTCCATCAGTTTGAAGTGATG 837
DB 799 IleAspAsnLeuSerSerPheLeuGlnSerGlyLeuAlaGlnThrValValSerValGly 818
QY 838 CTTCGACGTGGTGTGTTGTTATTAACAATGC-----GGTGGCCAGTTGCTGGTAAC 891
DB 819 ThrIleuIleGlyValValThrMetLeuAlaIleThrAspAlaGlnLeuAlaLeuValAla 838
QY 892 CTTCGAAACACTGTGTATACATACAGACATTCACAGTTGACGATCAGCGGTGAGAACTAGA 951
DB 839 LeuSerValValProIleIleIleValLeuThrLeuIlePheArgArgIleSerSerArg 858
QY 952 TTTGAATGAAATGAAACAAAGCAGATATATGATGACGTAATGCTGCTATGACTCAGTG 1011
DB 859 LeuThrThrAlaSerArgGlnGlnAlaSerGlnValAsnAlaValPheHisGlnSerIle 878
QY 1012 CTGAATTATGAACTGTGATATTTTAATGAAAGATATGAAAGACAGACATATGAT 1071
DB 879 AlaGlyLeuAlaGlnThrAlaGlnMetHisArgMetGlnSerGlnValPheAspAsnThrAla 898
QY 1072 GGATTTTGAAGCGTATGAGACTGCTTCATTTGAAAAGTACCTGCTGCTATGCTG 1131
```

```

DB 899 GlyGluAlaGlnGluPheArgArgLeuArgValIleSerGlnThrAlaIleAlaIleTyr 918
QY 1132 AACTTGTGCAAAAGTGTATTTTTCAGTGTGCTTTTAAACAGTAAATAGTGTCCAGT 1191
DB 919 PheProGlyLeuGlyAlaLeuSerGlnIleAlaGlnAlaLeuValLeuGlyPheGlyAla 938
QY 1192 CAGGAAATGTGTGCGAGTACCTTACTGTTGAGATCTA-----GTAAATGTGAAT 1242
DB 939 LeuGlnValThrArgGlyAspIleSerThrGlyValLeuValAlaPheValLeuTyrMet 958
QY 1243 GGAATGCTTTTTCAGCTTTTCAATTAACCCCTGAACTTTCTGGAACTGATATAGAGACT 1302
DB 959 GlyLeuMetPheGly-----ProIleGlnGlnLeuSerGlnIlePheAspSerTyr 975
QY 1303 AGACAAACCTCATGATATGAAACACCTGTTTACTCTACTCAAGTATGACACCAATAT 1362
DB 976 GlnGlnAlaAlaValAlaGlyPheArgArgIleThrGlnLeuLeuAlaThrGlnProSerVal 995
QY 1363 AAAGACAAAGTATGAGTATGCTCCCTTCAGATCACACACAGACAGCTACC-----GTG 1416
DB 996 ProAsp-----LeuGlyThrAsnArgHisAlaArgGlnAlaIleThrGlnProLeu 1012
QY 1417 GCCTTGATATATGATGCAATTTTGAATACATTGAGGGCCAGAAAGTCTTATGCAATATCC 1476
DB 1013 LeuLeuAspAspValThrPheGlyTyrSerAsp---AspProIleLeuAspAsnValThr 1031
QY 1477 TTGAACTCCCTGACGAAAGAAAGAAAGTCCATTTGTAGAGTATGGGTGACGAAAGC 1536
DB 1032 ValGlnIleGluProGlyThrThrValAlaValAlaGlyProThrGlyAlaGlySer 1051
QY 1537 ACAATAGTGAAGCTATATTTCCCTTCATAGAGCTCAAAAGGTAGCATTTATCTTGCT 1596
DB 1052 ThrValValIleLeuLeuSerArgLeuTyrAspProAsnGlnGlyAlaValIleGly 1071
QY 1597 GGTCAAAATATACAGATGTGAGCTGGAAGAGCCTTGGAGGCGAGTGGAGTGGTACT 1656
DB 1072 GlnIleAspIleLeuAspPheProThrAlaAspTrpArgArgThrIleGlyThrValPro 1091
QY 1657 CAGAGTCTGCTCCCTTCATTAATACATTTATTAACAACCTTATTTGAAACATCAGT 1716
DB 1092 GlnGlnAlaHisLeuPheSerGlySerIleAlaAspAsnIleGlyTyrGlyCysArgGlu 1111
QY 1717 GCTTCACTGAGGAGTGTGCAAGTGGCAAAATTAAGCTGACTTCAATGATGCAATCTT 1776
DB 1112 AlaSerThrSerIleGlyAlaAlaAlaArgArgValIleAlaLeuAlaIleAla 1131
QY 1777 GGAATGCCATGAGATATGACACCCAAAGTAGGGAACGAGACTCAAGCTTTGACAGAGA 1836
DB 1132 AlaIleProAspGlyPheAsnHisGlnValGlyGlnArgGlyArgAsnLeuSerSerGly 1151
QY 1837 GAAAGCAAAAGATGAGATATTCGAAGGCCATTTTGAAGAACCCCGACATATCTTAT 1896
DB 1152 GlnArgGlnLeuIleAlaLeuAlaArgAlaGlnLeuIleLeuProSerIleMetLeuLeu 1171
QY 1897 GATGAGCTACTTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1956
DB 1172 AspGlnAlaThrSerThrLeuAspProAlaThrGlnAlaValIleLeuAsnAlaSerAsp 1191
QY 1957 GATGTGTCAAAACAGCAAACTTATTTTTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2016
DB 1192 ArgValThrIleGlyArgThrSerIleIleValAlaHisArgLeuAlaThrAlaIleArg 1211
QY 2017 GCAGATGAAATCATGTCTGATCAGAGGTGAAGTACGACGACGACGACGACGACGACGACGAC 2076
DB 1212 AlaAspArgIleLeuValValGlnGlnGlyArgIleIleGlnAspGlySerHisAspAla 1231
QY 2077 TTGCTGCTAACCTCATATGATATCTATTCAGAAATGTGGAT 2118
DB 1232 LeuLeu---SerAlaAsnGlyThrTyrAlaArgMetTrpHis 1244

RESULT 15
US-09-873-409-6
```


Db 574 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 593

Search completed: February 15, 2003, 06:10:50
Job time : 40.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 05:54:00 ; Search time 20.5 Seconds
(without alignments)
6731.390 Million cell updates/sec

Title: AF133659
Perfect score: 4254
Sequence: 1 ATGGCCGCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table:
BLASTSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+n2p model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/AF133655/runat.10022003_155441.29551/app.query.fasta.1.2503
-DB=Issued_Patents_AA -OPMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NOEM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRA=AF133659 @CGN 1.1 18 @runat.10022003.155441.29551 -NCPU=6 -ICPU=3
-NO_XLXPY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-MAFN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB.pcp:*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pcp:*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pcp:*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pcp:*
5: /cgn2_6/ptodata/2/1aa/PCPUS COMB.pcp:*
6: /cgn2_6/ptodata/2/1aa/backfilest.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3809	89.5	747	2	US-08-895-522-1
2	3809	89.5	747	3	US-09-195-391-1
3	3342	78.6	694	2	US-08-895-522-3
4	3342	78.6	694	3	US-09-195-391-3
5	1516.5	35.6	694	2	US-08-895-522-4
6	1516.5	35.6	694	3	US-09-195-391-4
7	1201.5	28.2	475	4	US-09-212-247C-4
8	695.5	16.3	582	4	US-09-134-001C-4730
9	641.5	15.1	707	4	US-08-772-270A-4
10	637.5	15.0	580	4	US-09-134-001C-5611
11	618.5	14.5	711	3	US-08-772-270A-12
12	594.5	14.0	791	1	US-08-394-880B-2

13	583.5	13.7	1275	3	US-09-120-513-2	Sequence 2, App11
14	583.5	13.7	1275	4	US-09-450-105-2	Sequence 2, App11
15	583	13.7	1279	2	US-08-784-649A-2	Sequence 2, App11
16	578.5	13.6	1280	6	US-08-583-276-19	Sequence 19, App11
17	578.5	13.6	1280	6	5206352-4	Patent No. 5206352
18	565.5	13.3	1280	2	US-08-752-447-2	Sequence 2, App11
19	565.5	13.3	1280	4	US-09-316-167-2	Sequence 2, App11
20	562.5	13.2	1334	2	US-08-996-545-2	Sequence 2, App11
21	562.5	13.2	1334	4	US-09-328-320-2	Sequence 2, App11
22	562	13.2	717	4	US-08-924-629C-5	Sequence 5, App11
23	554	13.0	1408	1	US-08-612-521-2	Sequence 2, App11
24	546	12.8	1307	1	US-08-395-246C-2	Sequence 2, App11
25	545.5	12.8	1349	2	US-08-612-734B-2	Sequence 2, App11
26	536	12.6	1308	2	US-08-996-644-2	Sequence 2, App11
27	536	12.6	1308	3	US-09-352-552-2	Sequence 2, App11
28	518.5	12.2	748	4	US-09-061-764A-19	Sequence 19, App1
29	513	12.1	400	4	US-08-961-083-190	Sequence 190, App
30	509	12.0	1263	4	US-09-351-224E-11	Sequence 11, App1
31	507.5	11.9	1302	1	US-08-332-537-2	Sequence 2, App11
32	470.5	11.1	686	4	US-09-061-764A-15	Sequence 15, App1
33	442	10.4	1437	3	US-09-061-400-2	Sequence 2, App11
34	442	10.4	1453	2	US-09-001-273-2	Sequence 2, App11
35	442	10.4	1453	4	US-08-843-459A-2	Sequence 2, App11
36	428	10.1	1622	4	US-08-972-927-6	Sequence 6, App11
37	421	9.9	1621	4	US-08-972-927-3	Sequence 6, App11
38	414.5	9.7	653	4	US-09-061-764A-2	Sequence 2, App11
39	408	9.6	1261	4	US-09-605-785-538	Sequence 538, App
40	406	9.5	1261	4	US-09-439-313-538	Sequence 538, App
41	406	9.5	327	1	US-08-463-092B-9	Sequence 9, App11
42	406	9.5	327	2	US-08-460-907B-9	Sequence 9, App11
43	401.5	9.4	1531	1	US-08-141-893-2	Sequence 2, App11
44	401.5	9.4	1531	1	US-08-463-092B-2	Sequence 2, App11
45	401.5	9.4	1531	1	US-08-463-092B-4	Sequence 4, App11

ALIGNMENTS

RESULT 1
US-08-895-522-1
; Sequence 1, Application US/08895522
; Patent No. 5858719
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Puryi
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,522
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: Pf-0336 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 747 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: OVARNOT02
/ CLONE: 545981
/ US-08-895-522-1

Alignment Scores:
Pred. No.: 0 Length: 747
Score: 3809.00 Matches: 746
Percent Similarity: 99.87% Conservative: 0
Best Local Similarity: 99.87% Mismatches: 1
Query Match: 89.54% Indels: 0
DB: Gaps: 0

AF133659 (1-2345) x US-08-895-522-1 (1-747)

QY 16 ATGCATTCTGGCGCTGGCGGCGCGGCGCTCTTGGAAAAGCCGCGACCTCCGCG 75
DB 1 MetHisSerTPrArgTrpAlaAlaAlaAlaAlaPheGluLysArgHisSerAla 20
QY 76 ATTTCGATCCCGGCTTTAGTCTCTGTTAGCGGCTCAGGTCGCCAGTGGAGGCGACATCA 135
DB 21 IleLeuIleArgProLeuValSerValSerGlySerGlyProGlnTrpArgProHisGln 40
QY 136 CTGGCGCTTGGGAACCGCTCGAGCGCTACAGATTCAGAGTCTCAATTAAGTAATCA 195
DB 41 LeuGlyAlaLeuGlyThrAlaArgAlaTyrglnIleProGlnSerLeuLysSerIleThr 60
QY 196 TGGCAGAGATTGGAAAAGGCAATTCAGACAGTCTTAGATGCTGCAAAAGGCTCTCCAG 255
DB 61 TrpGlnArgLeuGlyLysGlyLysSerGlyGlnPheLeuAspAlaAlaLysAlaLeuGln 80
QY 256 GATATGGCCCTGATAGAAAAGAGACATGTGGCATGTGTCATGACGAGAGAGAGACTCCAC 315
DB 81 ValTrpProLeuIleGluLysArgThrCysTrpHisGlyHisAlaGlyGlyLysLeuHis 100
QY 316 ACAGACCCCAAGAAAGGTTAAAGATGTGATCTCGAAAATGATCAATAAGCAATGCTT 375
DB 101 ThrAspProLysGlnGlyLysLysAspValAspThrArgLysIleIleLysAlaLysLeu 120
QY 376 TCTTATGTGTGGCCCAAGACAGCGCCAGATCTACAGAGTGAAGTTGCCATTTGCGTGGGA 435
DB 121 SerTyrrValTrpProLysAspArgProAspLeuArgAlaArgValAlaIleSerLeuGly 140
QY 436 TTTTGGGTGGTGCAAAAGGCCCATGAATATGTGGTTCCTTCAATGTTTAAATATGCTGTA 495
DB 141 PheLeuGlyGlyAlaLysAlaMetCysnIleValIleProPheMetPheLysTyrrAlaVal 160
QY 496 GACAGCGCTCAACAGATGTGCGGAAAACATGTGAAGCTGAGATGACCAATAATCAAGTT 555
DB 161 AspSerLeuAsnGlnMetSerGlyAsnMetLeuAsnLeuSerAspAlaProAsnThrVal 180
QY 556 GCAACCATGCAACAGCAGTCTGATGTGGCTATGTGTTCAAGAGCTGAGACTGCTTTT 615
DB 181 AlaThrMetAlaThrAlaValIleuIleGlyTyrglyValSerArgAlaGlyAlaAlaPhe 200
QY 616 TTTTACGAAGTTGCAAAATCAGTATTTGGCAAGTACGCCAAGATTCATTCGAAGAATA 675
DB 201 PheAsnGlnValAlaArgAsnAlaValPheGlyLysValAlaGlnAsnSerIleArgArgIle 220
QY 676 GCCAAAATATGTTTCTCATCTTCACACCTGGAATCTGGGTTTCACTGAGCGACAG 735
DB 221 AlaLysAsnValAlaPheLeuHisLeuHisAsnLeuAspLeuGlyPheHisLeuSerArgGln 240
QY 736 ACGGAGGCTTTATCTAAGGCTATTGACAGAGAGAACAGAGGCTATCAGTTTGTCTGAGT 795
```

```
DB 241 ThrGlyAlaLeuSerLysAlaIleAspArgGlyThrArgGlyIleSerPheValLeuSer 260
QY 796 GCTTGGTATTTAATCTCTCCCATCATGTTTGAAGGATGCTGTGCAGTGCTTTG 855
DB 261 AlaLeuValPheAsnLeuLeuProIleMetPheGluValMetLeuValSerGlyValLeu 280
QY 856 TATTACAAATCCGGTCCCATGTTGCTTTGGTAAACCTTGGACACTTGTATACACA 915
DB 281 TyrrTyrrLysCysGlyAlaGlnPheAlaLeuValThrLeuGlyThrLeuGlyThr 300
QY 916 GCATTACAGATTGACATGCACACCGGTGAGAACTGATTTAGAAATAGAAATCAACAAACA 975
DB 301 AlaPheThrValAlaValAlaThrArgTrpArgThrArgPheArgIleGluMetAsnLysAla 320
QY 976 GATATATGATGAGGAATATGCGCTATAGACTCACCTGCGAATTAATGAACGTGAAGTAT 1035
DB 321 AspAsnAspAlaGlyAsnAlaAlaIleAspSerLeuLeuAsnTyrrGluThrValLysTyrr 340
QY 1036 TTTTATATGAAAAGATATGAAGCACAGAGATATGATGATTTTGAAGACGTATGAGACT 1095
DB 341 PheAsnAsnGlnLysArgTyrglyAlaIleGlnArgTyrrAspGlyPheLeuLysThrTyrrGluThr 360
QY 1096 GCTTCATTGAAAAGTACTCTTACTCTGTGCTATGCTGAACCTTTGGTCAAGTCTATTTTC 1155
DB 361 AlaSerLeuLysSerThrSerThrLeuAlaMetLeuAsnPheGlyGlnSerAlaIlePhe 380
QY 1156 AGTGCCTTTTACAGCATATATATGATGCTGCGCAGTCAAGGAAATTGGCAGTACCTT 1215
DB 381 SerValGlyLeuThrAlaIleMetValLeuAlaSerGlnGlyIleValAlaGlyThrLeu 400
QY 1216 ACTGTGGAGATCTAGTATATGATGATGAGACTGCTTTTTCAGCTTTCAATACCCCTGAAC 1275
DB 401 ThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeuProLeuAsn 420
QY 1276 TTTTGGGAACTGTATATATAGAGACTAGACACAGCTCATATGATATGAACCTTGT 1335
DB 421 PheLeuGlyThrValTyrrArgGluThrArgGlnAlaLeuIleAspMetAsnThrLeuPhe 440
QY 1336 ACTCTACTCAAGGATAGACACCCCAATTAAGCAAGATGATGCGATCCCTCCAGATC 1395
DB 441 ThrLeuLeuLysValAspThrGlnIleLysAspLysValMetAlaSerProLeuGlnIle 460
QY 1396 ACACACACAGACGCTACCGTGGCTTTGATATGATGCAATTTTGAATACATTGAGGGCCAG 1455
DB 461 ThrProGlnThrAlaThrValAlaPheAspAsnValHisPheGluTyrrIleGluGlyGln 480
QY 1456 AAAGTCTTAGTGAATATCTCTTGAAGTCCCTCGAGGAAAGAAAGTGCCATTGTAGA 1515
DB 481 LysValLeuSerGlyIleSerPheGluValProAlaGlyLysValAlaIleValAlaGly 500
QY 1516 GGTAGTGGGTGAGGAAAAGCAAAATAGTGAGGCTATTATTCGCTCTATGAGGCTCAA 1575
DB 501 GlySerGlySerGlyLysSerThrIleValArgLeuLeuPheArgPheTyrrGluProGln 520
QY 1576 AAGGTAGCATTTTATCTTGTGCTCAAAATATATACAGATGAGCTCGAAAGCCTTCGG 1635
DB 521 LysGlySerIleTyrrLeuAlaGlyGlnAsnIleGlnAspValSerLeuGlnSerLeuArg 540
QY 1636 AGGCGATGGGAGTGGTACCTCAGAGATGCTGCTCTTCCATTAATATCTATTATTAACAC 1695
DB 541 ArgAlaValGlyValValProGlnAspAlaValLeuPheHisAsnThrIleTyrrLysAsn 560
QY 1696 CTCTTATATGGAACATCAGAGCTTCACTCAGGAGAGGTATGGAGTGGCAAAATTTAGCT 1755
DB 561 LeuLeuTyrrGlyAsnIleSerAlaSerProGlnValTyrrAlaValAlaLysLeuAla 580
QY 1756 GGACTTCATGATGCAATCTTCGAATGCGCACATGATATGACACCCCAATGAGGGAACGA 1815
DB 581 GlyLeuHisAspAlaIleLeuArgMetProHisGlyTyrrAspThrGlnValGlyGluArg 600
QY 1816 GGACTCAAGCTTTTCAAGAGAGAAAGCAAAAGATAGCAATTTGACAGAGCCATTTTGAAG 1875
DB 601 GlyLeuLysLeuSerGlyGlyGlyLysGlnArgValAlaIleAlaArgAlaIleLeuLys 620
```


QY 1876 GACCCCCAGTACTCTATGATGAAGCTACTTCATGTTAGATTGCAATTCGAAG 1935
| | | | |
Db 621 AspProValIleLeuYAspGluAlaThrSerSerLeuAspSerIleThrGlu 640
| | | | |
QY 1936 ACTATCTGGTGGCAAGAGATGTGTCACACAGAACTTCTATTTTCATTCGCAC 1995
| | | | |
Db 641 ThrIleuGluAlaMetLysAspValValLysHisArgThrSerIlePheIleAlaHis 660
| | | | |
QY 1996 AGATTGTCACAGTGTGATGACAGTAAATCATTTGCTTGATCCAGGGTAAAGTACC 2055
| | | | |
Db 661 ArgLeuSerThrValValAspAlaAspGluIleIleValLeuAspGlnGlyValAla 680
| | | | |
QY 2056 GAACGTGGTACCAACCAATGTTTCTGTACACCTCATAGTATCTATTCAGAAATGTG 2115
| | | | |
Db 681 GluArgGlyThrHisGlyLeuLeuAlaAsnProHisSerIleTyrSerGluMetTyr 700
| | | | |
QY 2116 CATTCACAGACAGCCGTGTGCAGACACATATTAACCCCAATGGAGCAAGAAAGAA 2175
| | | | |
Db 701 HisThrGlnSerSerArgValGlnAsnHisAspAsnProLysTyrGluAlaLysGlu 720
| | | | |
QY 2176 AATATATCCAAAGAGAGAAAGAAAGAACTACAGAAATTTGCAATAGTGTGAA 2235
| | | | |
Db 721 AsnIleSerLysGluGluArgLysLysLeuGlnGluIleValAsnSerValLys 740
| | | | |
QY 2236 GGCTGTGAAACTGTGTGTGC 2256
| | | | |
Db 741 GlySerGlyAsnCysSerCys 747
| | | | |
RESULT 2
US-09-195-391-1
Sequence 1, Application US/09195391
Patent No. 6080842
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Puri, Neil C.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
TITLE OF INVENTION: TRANSPORT PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09195,391
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/895,522
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0336 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: OVARNOT02
CLONE: 545981
US-09-195-391-1
Alignment Scores:
Pred. No.: 0
Score: 3809.00
Percent Similarity: 99.87%
Best Local Similarity: 99.87%
Query Match: 89.54%
DB: 3
Length: 747
Matches: 746
Conservative: 0
Mismatch: 1
Indels: 0
Gaps: 0
AF133659 (1-2345) x US-09-195-391-1 (1-747)
QY 16 ATGCAATCTTGGGCGCGGCGCGCGCGCGCTTTCGAAAAGCGCGCACTCCGCG 75
| | | | |
Db 1 MetHisSerTyrArgTrrAlaAlaAlaAlaAlaAlaPheGluLysArgArgHisSerAla 20
| | | | |
QY 76 ATTCTGATCCGCGCTTGTAGTCTCTGTAGCGGCTCAGGTCGCGAGTGCAGGCACATCAA 135
| | | | |
Db 21 IleLeuIleArgProLeuValSerValSerGlySerGlyProGlnTrrArgProHisGln 40
| | | | |
QY 136 CTGGCGCGCTTGGGAAACCGCTCGAGCCTTACAGATTCAGAGTCATTAAGATATACA 195
| | | | |
Db 41 LeuGlyAlaLeuGlyThrAlaArgAlaTyrGlnIleProGlnSerLeuLysSerIleThr 60
| | | | |
QY 196 TGGCAGAGATTTGGGAAAAGGCAATTCAGGACAGTCTTGTAGTCTGCGAAAGCGCTCCAG 255
| | | | |
Db 61 TrpGlnArgLeuGlyLysGlyAsnSerGlyGlnPheLeuAspAlaAlaLysAlaLeuGln 80
| | | | |
QY 256 GTATGGCACTGATTAAGAAAGAGACATGTGGCATGTGTCATGCAGAGAGAGACTTCAC 315
| | | | |
Db 81 ValTrrProLeuIleGluLysArgGlnCysTrrHisGlyHisAlaGlyGlyLeuHis 100
| | | | |
QY 316 ACAAGCCCAAAAGAGAGGCTTAAAGATGTTGATCTGCGAAATTCATAAAGCAATGCTT 375
| | | | |
Db 101 ThrAspProLysGluGlyLeuLysAspValAspThrArgLysIleIleLysAlaLysLeu 120
| | | | |
QY 376 TCTTATGTGGGCGCAAGAGCAGGCGCATCTCAGAGCTTACAGTTCGATTTGGCTGGGA 435
| | | | |
Db 121 SerTyrValTrrProLysAspArgProAspLeuArgAlaArgAlaAlaIleSerLeuGly 140
| | | | |
QY 436 TTTTGGGTGGTCAAAAGGCATGAATATGTGTGCTCCCTTCATGTTTAATATGCTGTA 495
| | | | |
Db 141 PheLeuGlyGlyAlaLysAlaMetAsnIleValValProPheMetPheLysTyrAlaVal 160
| | | | |
QY 496 GACAGCCTCAACCAAGATGTGGGAAAATGCTGAGCTGATGATGACCAAAATACATT 555
| | | | |
Db 161 AspSerLeuAsnGlnMetSerGlyAsnMetLeuAsnLeuSerAspAlaProAsnThrVal 180
| | | | |
QY 556 GCACACATGCGAACACAGATGTCGATTTGGCTATGTGTATCAAGAGCTGAGCTGCTTTT 615
| | | | |
Db 181 AlaThrMetAlaThrAlaValLeuIleGlyTyrGlyValSerTrrArgAlaGlyAlaAlaPhe 200
| | | | |
QY 616 TTTACGAAGTTGGAATGAGATTTGGCAAGTGGAGCCAGAAATTCATCCGAAGATA 675
| | | | |
Db 201 PheAsnGlyValArgAsnAlaValPheGlyLysValAlaGlnAsnSerIleArgGlnIle 220
| | | | |
QY 676 GCCAAAATGTCTTCTCATCTTCAAACTTGATCTGGGTTTTCACCTGACAGACAG 735
| | | | |
Db 221 AlaLysAsnValPheLeuHisLeuHisAsnLeuAspLeuGlyPheHisLeuSerArgGln 240
| | | | |
QY 736 ACGGAGCTTATCTTAAGCTATTTGACAGAGAAACAAGGGTATTCAGTTTGTCTGAGT 795
| | | | |
Db 241 ThrGlyAlaLeuSerLysAlaIleAspArgGlyThrArgGlyIleSerPheValLeuSer 260
| | | | |
QY 796 GCTTGTATTTATCTTCTTCCATCATGTTGAAGTAGTGTTCAGTGGTGTGTTT 855
| | | | |
Db 261 AlaLeuValPheAsnLeuLeuProIleMetPheGlnValMetLeuValSerGlyValLeu 280
| | | | |
QY 856 TATTACAAATGCGGTCAGATTTGCTTTGTAACTTGGAACTTGGTACATACACA 915
| | | | |

Db 281 TyrTyrLysCyseGlyAlaGlnPheValIleuValThrLeuGlyThrLeuGlyThrTyrThr 300
 Qy 916 GGATTCACAGTGGACGTGCACAGGTGAGAACTAGATTGAGATGAATAAGCAAGCA 975
 Db 301 AlaPheThrValAlaValAlaThrArgThrArgThrArgPheArgIleGluMetCysnysAla 320
 Qy 976 GATTAATGATGACAGTAATGCTGCTATAGACTGACTGCTGGAATTAATGAACCTGTGAAGTAT 1035
 Db 321 AspaAsnApAlaGlyAsnAlaAlaIleAspSerLeuLeuAsnTyrGluThrValLysTyr 340
 Qy 1036 TTTATATATGAAGAATATGACACACAGATATGATGATTTTGAAGAAGCTATGAGACT 1095
 Db 341 PheAsnAsnGluArgTyrGluAlaGlnArgTyrAspGlyPheLeuLysThrTyrGluThr 360
 Qy 1096 GCTTATGTAAGAAAGTACCTCTACTGCTGCTGATGTAACCTTGGTCAAAAGTCTATTTTC 1155
 Db 361 AlaSerLeuLysSerThrSerThrLeuAlaMetLeuAsnPheGlyGlnSerAlaIlePhe 380
 Qy 1156 AGTGTGGTTTAAACAGCTATTAATGGTGTGCTGCGCAGTCAAGGAATTTGTGCAGTACCTT 1215
 Db 381 SerValGlyLeuThrAlaIleMetValLeuAlaSerGlnGlyIleValAlaGlyThrLeu 400
 Qy 1216 ACTGTGGAGATCTTGTATATGCTGATGATGACTGCTTTTACGCTTTCATTACCTCGTAAC 1275
 Db 401 ThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeuProLeuAsn 420
 Qy 1276 TTTCTGGGGAACGTATATATGAGAGACTAGACAAGCACTGATGATATGAACCTTGTTT 1335
 Db 421 PheLeuGlyThrValTyrArgGluThrArgGlnAlaLeuIleAspPheCysnThrLeuPhe 440
 Qy 1336 ACTTACTCAAGGTAGACACCCCAATTAAGAACAAGTGATGGCACTCTCCCTTGAGATC 1395
 Db 441 ThrLeuLeuLysValAspThrGlnIleLysAspLysValMetAlaSerProLeuGlnIle 460
 Qy 1396 ACAACAACAGACGCTACCTGCTGCTTGTATATGTCATTTTGAATPACATTAGAGCCAC 1455
 Db 461 ThrProGlnThrAlaThrValAlaPheAspAsnValIleAspGluTyrIleGluGln 480
 Qy 1456 AAAGTCTTAGTGAATATCTTTGAAGTCCCTGACAGGAAAGAAAGTGCCATTGAGGA 1515
 Db 481 LysValIleSerGlyIleSerPheGluValProAlaGlyLysValAlaIleValGly 500
 Qy 1516 GGTACTGGGTGAGGAGAAAGCAAACTAGTAGGCTATTATTTGCTTATGAGGCTCA 1575
 Db 501 GlySerGlySerGlyLysSerThrIleValArgLeuLeuPheArgPheTyrGluProGln 520
 Qy 1576 AAGGTAGCATTTATCTGCTGCTGCAAAATATACAAAGATGAGGCTGGAAGGCTTGG 1635
 Db 521 LysGlySerIleTyrLeuAlaGlyGlnAsnIleGlnAspValSerLeuGluSerLeuArg 540
 Qy 1636 AGGGAGTGGAGTGTACCTCAGAGTGTCTCTTCATATATCTATTATTAACAAC 1695
 Db 541 ArgAlaValAlaGlyValValProGlnAspAlaValLeuPheAsnThrIleTyrTyrAsn 560
 Qy 1696 CTCTTATATGGAACATCAGTCTTCACTGAGGAAGTATGACAGTGCAGCAAAATTAAGT 1755
 Db 561 LeuLeuTyrGlyAsnIleSerIleAspProGluGluValTyrAlaValAlaLysLeuAla 580
 Qy 1756 GGACTTCATGATGCAATTTTGAATGCCACATGATATGACACCCAGTAGGGGAACGA 1815
 Db 581 GlyLeuHisAspAlaIleLeuArgMetProHisGlyTyrAspThrGlnValGlyLysArg 600
 Qy 1816 GGAGCTCAAGCTTTCAGGAGGAGAAAGCAAGATGAGCAATTCGAAGCCATTTGAG 1875
 Db 601 GlyLeuLysLeuSerGlyGlyLysGlnArgValAlaIleAlaArgAlaIleLeuLys 620
 Qy 1876 GACCCCAAGTCATCTCTATGATGAGAGTACTTATCTGATTCGATTACTGAAG 1935
 Db 621 AspProPheValIleLeuLysArgIleAlaThrSerSerLeuAspSerIleThrGluGlu 640
 Qy 1936 ACTATTTTGGTGCATGAGAGATGTGTCAAACACAGAACTTCTATTGTCACAC 1995
 Db 641 ThrIleLeuGlyAlaMetLysAspValValLysHisArgThrSerIlePheIleAlaHis 660

Qy 1996 AGATTGTCACAGTGGTGTGATGACAGTGAATCATGCTTTGGATCAGGGTAAAGTACCC 2055
 Db 661 ArgLeuSerThrValValAspAlaAspGluIleIleValLeuAspGlnGlyLysValAla 680
 Qy 2056 GAACGTGTATCCACCATGCTGTTGCTTGTCTTACCTCATAGTATCTTTCAGAAATGTG 2115
 Db 681 GluArgGlyThrHisIleGlyLeuLeuAlaAsnProHisSerIleTyrSerGluMetTyr 700
 Qy 2116 CATACACAGACACCGGTGTCAGAACCATGATTAACCCCAATGGAAGCAAAAGAA 2175
 Db 701 HisThrGlnSerSerArgValGlnAsnHisAspAsnProLysTrpGluAlaLysLysGlu 720
 Qy 2176 AATATATCAAGAGAGAGAAAGAAAGAACTACAGAGAAATTTGCAATAGTGTGAA 2235
 Db 721 AsnIleSerLysGluGluGluArgLysLysLeuGlnGluIleValAlaSerValLys 740
 Qy 2236 GGCTGTGAAACTGTTGTGTC 2256
 Db 741 GlyCysGlyAsnCysSerCys 747

RESULT 3

US-08-895-522-3
 ; Sequence 3, Application US/08895522
 ; Patent No. 5858719
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/895,522
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0336 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 694 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: Genbank
 ; CLONE: 1167982
 ; US-08-895-522-3

Alignment Scores:

Pred. No.: 0 Length: 694
 Score: 3342.00 Matches: 656

```

AF133659 (1-2345) x US-08-895-522-3 (1-694)
Percent Similarity: 97.41% Conservative: 20
Best Local Similarity: 94.52% Mismatches: 18
Query Match: 78.56% Indels: 0
DB: 2 Gaps: 0

```

QY	175	GAGCATTTAAAAAGATACACATGCGCAGAGANTGGGAAAAGCAATTCAGACACAGTTCTTA	234
Db	1	GlubserleuAsganInmThrGlnGlnThrGTrpGlyAspSbnSerArgInIleuLeu	20
QY	235	GATGCTGCAAAAGGCTCTCCAGGATAGCGCCTGATAGAAAAGGACATGTTGGCATGCT	294
Db	21	AspIatThrIysAlaIeuGlnInThrTrpProIeuIleGlyIuAsrGlnCysTrpHisGly	40
QY	295	CATCAGAGAGAGACCTCCACACAGACCCAAAAGAGGTTTAAAGATGTTGATCTCGG	354
Db	41	HisIaGlyGlyGlyLeuHisIsthAspProIyGeliuGlyLeuSAspValaAspThrArg	60
QY	355	AAATCATTAATAAGCAAAAGCTTTCTTATGTGTGGCGCCCAAAACAGAGGCAGATCTACAGCT	414
Db	61	LysIleIleuYsAlaMetIleuSerTyValTrpProGluAspArgProAspLeuArgAla	80
QY	415	AGAGTTGCCATTTGCTGGGATTTTGGGTGGTGCAAAAGGCATGAATATTGGTTCCC	474
Db	81	ArgValAlaIleSerIeuGlyPheIeuGlyGlyAlaIysAlaIysAlaMetAsnIleValPro	100
QY	475	TTCAAGTTTAAATATGCGTGTGAACGGCTCAACCAAGATGTCGGGAAACATCTGAAACCTG	534
Db	101	PheMetPheIysTrpAlaValAspSerIeuSbnGlnMetSerGlyAsnMetLeuAsnIleu	120
QY	535	AGTGATGCACCAATATACAGTTGCACACATGGCAACAGACAGTTCTGATGTGCTATGGTGA	594
Db	121	SerAspAlaProAsnThrValAlaThrMetAlaThrAlaValIleuIleGlyTrpGlyVal	140
QY	595	TTCAGAGCTGGAGCTGCTTTTTHAAGAAAGTTGAAATGCAGTATTGGCAGAGTACGCC	654
Db	141	SerArgAlaGlyAlaAlaPhePheAsnGlyValArgSbnAlaValPheGlyLysValAla	160
QY	655	CAGATTCATCCCAACAAATAATGACCAAAAAGTCTTTCATCATCTTCCACATCGATCTCG	714
Db	161	GlnSbnSerIleArgGlyIleAlaIysAsnValPheIeuHisIleuHisSbnIleuAspLeu	180
QY	715	GGTTTCACTGAGCAGACACAGCGGAGCTTTATCTAAGCTATTGACAGAGAACAGG	774
Db	181	GlyPheHisIleuSerArgSbnGlnThrGlyAlaIeuSerIysAlaIleAspArgGlyThrArg	200
QY	775	GGATACGTTTGGCTCGAGAGCTTGGATTTAATATCTTCCCAATCAGTTTGAAGG	834
Db	201	GlyIleSerPheValIleuSerAlaIeuValPheSbnIleuProIleValPheGluMet	220
QY	835	ATGCTTTCAGTGGTGTTTTGATTTTACAAATGCGGTGCCAGTTGTGGTAACCTT	894
Db	221	MetIeuValSerSerValIeuTrpTyLysCysGlyAlaGlnPheAlaIeuValThrIleu	240
QY	895	GGACACCTTGTAATACACAGCATTTACAGTTGCCAGTCAACGGTGTGAGAAACTAAGATT	954
Db	241	GlyThrIeuGlyAlaTrpThnAlaPheThrValAlaValaThrArgTrpArgThrArgPhe	260
QY	955	AGAAATAGAAATGAACAAGCAGATAAAGATGACAGGTAATGCTCTATAGACTCACTGCTG	1014
Db	261	ArgIleGluMetAsnIysAlaAspSbnAspAlaGlyAsnAlaAlaIleAspSerIleuLeu	280
QY	1015	AATTATGAACCTGTGAAGATATTTTAAATGAAGAAGATATAGACACAGAGATATGATGA	1074
Db	281	AsnTrpGlnThrValLysTrpPheAsnAsnGluIysTrpGluAlaGlnArgTrpAspGly	300
QY	1075	TTTTTGAAGACGTATGAGACGCGCTTCAATTGAAAGATCCCTTACTCGGTATGCTGAAC	1134
Db	301	PheIeuIysThrTrpGluThrAlaSerIleuIysSerThrSerThrIleuAlaMetIleuAsn	320
QY	1135	TTTGCTCAAAAGTGTATTTTTCAGTGTGCGTTTAAACAGCTATTAATGTGTCTGCGCACTCAG	1194

Db	321	phegilyglnasna1a1lepheserValgilyleuthrAla1leMeValleuAlaserGln	340
Qy	1195	GGAAATGGGCGAGTACCCTTAAGTGGAGATCTAGTAATGGTAATGAGACTGCTTTT	1254
Db	341	GlyIleValAlaGlyAlaLeuthrValGlyAspLeuValMetValAsnGlyLeuPhe	360
Qy	1255	CAGCTTTCATTACCCCTGAACTTTCTGGGAACGTATATATAGAGACTATGACAAAGACTTC	1314
Db	361	GlnLeuSerLeuProLeuAsnPhLeuGlyThrVal1TyArgGlyuthrArgGlnAlaLeu	380
Qy	1315	ATAGATATGAACACCTTGTTTACTGTACTCAAGGTAGACACCCCAAAATTPAAACAAGTG	1374
Db	381	IleAspMetAsnThrLeuPheThrLeuLeuValValAspThrArg1LeuAspIlyVal	400
Qy	1375	ATGGCATCTCCCTTCAGATCAACACAGACAGCTACCGCTGGCTTGTATATGTGCAT	1434
Db	401	MetAlaProProLeuGlnIleThrProGlnThrAlaThrValAlaPheAspAsnValHis	420
Qy	1435	TTTGAATACATTGAGGGCGAGAAAGTCTTACTGTGGAATATCTTTGAAAGTCCCTGCAGGA	1494
Db	421	PheGlyTyrIleGlnGlyAlaLeuAsnGlyValLeuAsnGlyValSerPheGlnValProAlaGly	440
Qy	1495	AAGAAAGGGCCATTGTAGAGGTGTGGGTGACAGGAAAGAAAGCAATGTGAGGCGATTA	1554
Db	441	LysIlyValAlaIleValAlaGlyIlySerIlySerGlyIlySerThrIleValArgLeuLeu	460
Qy	1555	TTTCCTCTTATGAGCCCTCAAAAGGGTAGCATTTATCTTGCTGTCAAAATATACAGAT	1614
Db	461	PheArgPheTyrGlnProGlnIlySerIleTyrIleAlaIaGlyGlnAsnLeuGlnAsp	480
Qy	1615	GTGAGCTTGAAAGCTTCGGAGGGCAGTGGAGTGTACTCAGAGATGCTGCTCTTC	1674
Db	481	ValSerLeuGlnSerLeuArgArgAlaValAlaGlyValAlaProGlnAspAlaValLeuPhe	500
Qy	1675	CATAATCTATTTATTAACAACCTCTTATATGAAACATCAAGTCTTACCTGAGGAAGTG	1734
Db	501	HisAsnThrIleTyrTyrAsnLeuLeuTyrGlyAsnIleAsnAlaSerProGlnIlyVal	520
Qy	1735	TATGAGTGGCAAAATTTAGCTGTGACTCATATGGAATTTCTTCAATCCCAATGAGATAT	1794
Db	521	TyrAlaValAlaIlyLeuAlaIaGlyLeuHisAspAlaIleLeuArgMetProHisGlyTyr	540
Qy	1795	GACACCCAAAGTGGGGAACGAGAACTCAAGCTTTGAGAGAGAAAGCAAAAGAGTACGA	1854
Db	541	AspThrGlnValAlaGlyIlyArgIlyLeuIlySerLeuSerGlyIlyGlnIlyGlnArgValAla	560
Qy	1855	ATTGCAAGAGCCATTTTGAAGACCCCCAGCTCATACTGTATGAAGACTTCTCATATG	1914
Db	561	IleAlaArgAlaIleLeuIlyAsnProValIleLeuTyrAspGlnAlaThrSerSer	580
Qy	1915	TTAGATGGATTACTGAGAGAGACTTCTTGGTGCCATGAAGAGATGGTGCACACAGA	1974
Db	581	LeuAspSerIleThrGlnGlnThrIleLeuGlyAlaMetAspAspValValIlyHisAspG	600
Qy	1975	ACTTCTATTTTCATGACACACAGATTGTCAACAGTGGTGTGATGACATGAATCATTTGTC	2034
Db	601	ThrSerIlePheIleAlaHisAspGlnSerThrValValAspAlaAspGlnIleIleVal	620
Qy	2035	TTTGATCAGGGGTAAAGTACCGGAACGTGTACCCCAACCATGGTTGCTTGTCTACCTCAT	2094
Db	621	LeuSerGlnGlyIlyValAlaGlyIlyArgIlyThrHisTyrGlyLeuLeuAlaAsnSerSer	640
Qy	2095	AGTATCTTATCAGAAATGGGATGATCAGAGAGACCGCTGGACAGAACCATATTAACCC	2154
Db	641	SerIleTyrSerGlnMetTyrHisIleThrGlnSerAsnArgValGlnAsnGlnAspSerLeu	660
Qy	2155	AAATGGAGAGCAAGAAAGAAATATATATCAAGAGAGAGAAAGAAAGAACTTACAGAA	2214
Db	661	GlyTyrAspAlaIlyIlyIlyGlnIlySerLeuSerIlyGlnGlnIlyArgIlyIlyLeuGlnGln	680
Qy	2215	GAATTTGCAATAGTGTGAAGGCTGTGGAACCTGTTGCTGC	2256
Db	681	GlnIleValAsnSerValIlyGlyIlyCysGlyIlyAsnCysSerCys	694

```

RESULT 4
US-09-195-391-3
; Sequence 3, Application US/09195391
; Patent No. 6080842
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,391
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/895,522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0336 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1167982
; US-09-195-391-3

Alignment Scores:
Pred. No.: 0 Length: 694
Score: 3342.00 Matches: 656
Percent Similarity: 97.41% Conservative: 20
Best Local Similarity: 94.52% Mismatches: 18
Query Match: 78.56% Indels: 0
Gaps: 0

AF133659 (1-2345) x US-09-195-391-3 (1-694)
QY 175 GAGTCATTAAAGATGATCATGACGAGATTGGAAAAAGCATTCAGACAGTCTTCA 234
Db 1 GlnSerLeuArgAsnThrThrlnglnlnrGtprGlyLysAspAsnSerAArglnleu 20
QY 235 GATGCTGCAAGGCTCTCCAGGTATGGCCACTGATTAAGAAAAGAGACATGTTGGCATGGT 294
Db 21 AspaLatrLysAlaLeuGlnThrTrpProLeuIleGlnLysArgThrCysTrpHisGly 40
QY 295 CATGAGAGAGAGAGACTCCACACAGACCAAGAAAGGTTAAAGATGGTATGATCTGG 354
Db 41 HisAlaGlyGlyGlyLeuHisThrAspProLysGlnGlyLeuLysAspAlaAspThrArg 60
QY 355 AAAATCATAAAGCAATGCTTCTTATGTGTGTGCCCAAGACAGCCAGATCTACGAGCT 414

```

```

Db 61 LysIleLeuLysAlaMetLeuSerThrValTrpProGlnAspArgProAspLeuArgAla 80
QY 415 AGAGTTGCAATTTGGCTGGGATTTTGGGTGGTCAAGGCAGATATATTGGTCTCCC 474
Db 81 ArgValAlaIleSerLeuGlyPheLeuGlyGlyAlaLysAlaMetAsnIleValValPro 100
QY 475 TTCATGTTTAAATATGCTGTAGACAGCTCCAAACAGATGTGGGAAACATGCTGAACCTG 534
Db 101 PheMetPheLysThrAlaValAspSerLeuAsnGlnMetSerGlyAsnMetLeuAsnLeu 120
QY 535 AGTATGACCAACCAATATACATTGCAACCATGGCAACACAGCTTGTGATGGCTATGGTGA 594
Db 121 SerAspAlaProAsnThrValAlaThrMetAlaThrAlaValLeuIleGlyThrGlyVal 140
QY 595 TCAAGAGCTGAGCTGCTTTTAAAGAGTTGGAATGAGTATTTGGCAAGTGCC 654
Db 141 SerArgAlaGlyAlaAlaPhePheAsnGlnValArgAsnAlaValPheGlyLysValAla 160
QY 655 CAGAAATTCATCCGAGAAATAGCAAAATGCTCTTCCATCTTCACAACTGATCTG 714
Db 161 GlnAsnSerIleArgAlaGlyIleAlaLysAsnValPheLeuHisLeuHisAsnLeuAspLeu 180
QY 715 GGTTCACCTGACGACACAGACGGAGCTTATCTAAGCTATTGACAGAGAACAGG 774
Db 181 GlyPheHisLeuSerArgGlnThrGlyAlaLeuSerLysAlaIleAspArgGlyThrArg 200
QY 775 GGTAATGCTTTTTCCTGAGAGCTTTGGTATTTAATCTCTCCATCATGATTTGAAGTG 834
Db 201 GlyIleSerPheValLeuSerAlaLeuValPheAsnLeuLeuProIleValPheGlnMet 220
QY 835 ATGCTTGTCAGTGTGTTTGTATTAACAAATGCGGTGCCAGTTGGTGTGAACCTT 894
Db 221 MetLeuValSerSerValLeuThrTyrllySecGlyAlaGlnPheAlaLeuValThrLeu 240
QY 895 GGAACACTGTGTACATACACAGATTCAACGTTGCACTCACCGGTGAGAACTAGATT 954
Db 241 GlyThrLeuGlyAlaThrAlaPheThrValAlaValAlaThrArgTrpArgThrArgPhe 260
QY 955 AGAATGAAGAAAGCAAGCAAGATGATGACAGTAAAGTGTGATTAAGTCACTG 1014
Db 261 ArgIleGlnMetAsnLysAlaAspAsnAspAlaGlyAsnAlaAlaIleAspSerLeuLeu 280
QY 1015 AATTATGAACCTGTGAAGATATTTAATGAAGATATGAAGACAGATATGATGA 1074
Db 281 AsnThrGlnThrValLysThrPheAsnAsnGlnLysThrGlnAlaGlnArgTyAspGly 300
QY 1075 TTTTGAAGACGATGAGACTGCTTCATGAAGAAAGTACTTACTGTGGCTAGCTGAAC 1134
Db 301 PheLeuLysThrTyrlGlnThrAlaSerLeuLysSerThrSerThrLeuAlaMetLeuAsn 320
QY 1135 TTTGCTAAATGCTATTTTCAAGTGGCTTAAACAGCTAATAGTGGTCCGAGTCAG 1194
Db 321 PheGlyGlnAlaAlaIlePheSerValGlyLeuThrAlaIleMetValLeuAlaSerGln 340
QY 1195 GGAATGTGGAGGATACCTTACTGTTGAGATCTAGTAATGTGAATGCACTGCTTTT 1254
Db 341 GlyIleValAlaGlyAlaLeuThrValGlyAspLeuValMetValAsnGlyLeuLeuPhe 360
QY 1255 CAGCTTTCATTAACCTCGAATCTTCTGGGAACTGTATATAGAGAGACTAGACAAGCTC 1314
Db 361 GlnLeuSerLeuProLeuAsnPheLeuGlyThrValTyrlArgGlnThrArgGlnAlaLeu 380
QY 1315 ATAGATATGAACACCTTGTACTCTACTCAAGGTAAACCCCAATTAAGACAAAGTG 1374
Db 381 IleAspMetAsnThrLeuPheThrLeuLeuLysValAspThrArgIleLysAspLysVal 400
QY 1375 ATGGCATCTCCCTTCAGATCACACACAGACAGTACCGGTGGCTTGAATATGTCAT 1434
Db 401 MetAlaProProLeuGlnIleThrProGlnThrAlaThrValAlaPheAsnValHis 420
QY 1435 TTTGAATACATTTGAGGGCAGAAAGTCTTATGTAATATCTTTGAAGTCCCTGCAAGA 1494

```

Db 421 PheGluTrpIleGluGlnGlnValLeuAsnGlnValSerPheGluValProAlaGly 440
QY 1495 AAGAAAGTGGCCATTGAGAGAGTAGGGGTCAAGGAAACACAAATAGTAGGCTATTA 1554
Db 441 LysLeuValAlaIleValGlnGlySerGlySerGlySerGlySerThrIleValArgLeu 460
QY 1555 TTTCGCTTCATAGAGCTCAAAAGGTAAGCTTTATCTGTGGTCAAAATATCAAGAT 1614
Db 461 PheArgPheTrpGluProGlnLysGlySerIleTrpLeuAlaGlnAsnLeuGlnAsp 480
QY 1615 GTGAGCTGGAAAGCTTCGAGAGGAGTAGGGTACCTGAGATGCTGCTCTTC 1674
Db 481 ValSerLeuGlnSerLeuAlaGlnValAlaGlnValAlaProGlnAspAlaValLeuPhe 500
QY 1675 CATATACTATTATTAACAACCTTTATATGAGAAACATCAAGTCTTCACTGAGAAATG 1734
Db 501 HisAsnThrIleTrpTrpAsnLeuLeuTrpGlyAsnIleAsnAlaSerProGlnGlnVal 520
QY 1735 TATGCAATGGCAAAATTAAGTGGACCTTCATGATGCAATTCTTGCAATGCCCATGATAT 1794
Db 521 TyrAlaValAlaLysLeuAlaGlnLysLeuHisAspAlaIleLeuArgMetProHisGlyTyr 540
QY 1795 GACACCCAGTAGGGGGAACGAGCACTCAAGCTTCAGAGAGGAAAGCAAAAGTAGTGA 1854
Db 541 AspThrGlnValGlnGlnArgGlyLeuLysLeuSerGlyGlnGlyLysGlnArgValAla 560
QY 1855 ATTGCAAGAGCATTTTGAAGAGACCCCACTCATCTATGATGAGTCACTTCACTG 1914
Db 561 IleAlaArgAlaIleLeuLysAsnProProValIleLeuTrpArgGlnAlaThrSerSer 580
QY 1915 TTAGATTGATTAAGAGACTATTTCTGGTCCATGAAAGATGGTGTCAAAACAGAGA 1974
Db 581 LeuAspSerIleThrGlnGlnThrIleLeuGlyAlaMetArgAspValIleLysHisArg 600
QY 1975 ACTTCTATTTTCATGAGACAGATTGTCAACAGTGGTGAAGAGTGAATGATTTGTC 2034
Db 601 ThrSerIlePheIleAlaHisArgLeuSerThrValAlaAspAlaAspGlnIleLeuVal 620
QY 2035 TTGGATCAGGGTAAAGTAGGCGAAGCTGTGTAACCACTGATGGTTGTGCTAAACCTCAT 2094
Db 621 LeuSerGlnGlnLysValAlaGlnArgGlyThrHisIleTrpGlyLeuLeuAlaAsnSerSer 640
QY 2095 AGTATCTATTCAGAAATGTGGCATATCAAGAGACGCGTGTGCAAGAACATGATTAACCCC 2154
Db 641 SerIleTrpSerGlnMetTrpHisThrGlnSerAsnArgValGlnAsnGlnAspSerLeu 660
QY 2155 AATGGAGACCAAGAAAGAAATATATCCAAAGAGAGGAAAGAAAGAAATCTCAAGAA 2214
Db 661 GlnTrpAspAlaLysLysGlnSerLeuSerLysGlnGlnArgLysLysLeuGlnGln 680
QY 2215 GAAATGTCAATAGTGAAGAGCTGTGAAACTGTTCGTGC 2256
Db 681 GlnIleValAsnSerValLysGlyCysGlyAsnCysSerCys 694

RESULT 5
US-08-895-522-4
; Sequence 4, Application US/08895522
; Patent No. 5858719
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,522
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0336 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 575393
US-08-895-522-4

Alignment Scores:
Pred. No.: 4,17e-147 Length: 694
Score: 1516.50 Matches: 304
Percent Similarity: 67.83% Conservative: 122
Best Local Similarity: 48.41% Mismatches: 169
Query Match: 35.65% Indels: 33
DB: 2 Gaps: 7

AF133659 (1-2345) x US-08-895-522-4 (1-694)

QY 355 AATAATCAATAAACAATGCTTTCTTATGATGTCGCCAAAGACAGCCAGATCTACAGACT 414
Db 94 LysIleLeuLysAspLeuPheArgTyrIleTrpProLysGlyAsnAsnLysValArgIle 113
QY 415 AGAGTTGCGCATTTCCGCTGGGATTTTGGGTGGTGAAGGCGCATGATATTGTTGCC 474
Db 114 ArgValLeuIleAlaLeuGlnLeuIleSerAlaLysIleLeuAsnValGlnValPro 133
QY 475 TTGATGTTAAATATATGCTGTAGACAGCTTCACAGATGTGGGAAACATGCTGAACCTG 534
Db 134 PhePhePheLysGlnThrIleAspSer-----MetAsnIle 145
QY 535 AGTGAGCACCAAAATACAGTTGCAACCATGGCAACA-----GCAAGTTCGATTGGCTAT 588
Db 146 AlaTrpAspAspProThrValAlaLeuProAlaAlaIleGlyLeuThrIleLeuCysTyr 165
QY 589 CGGTATCAAGAGCTGAGCTGCTTTTAAAGCAAGTCGAAATGAGATTTGGCAAG 648
Db 166 GlyValAlaArgPheGlySerValLeuPheGlyGlnLeuArgAsnAlaValPheAlaLys 185
QY 649 GTAGCCAGAAATTCAGAGATCGAAGATAGCCAAATATGCTTTCTTCATCTTCAACACTG 708
Db 186 ValAlaGlnAsnAlaIleArgThrValSerLeuGlnThrPheGlnHisLeuMetLysLeu 205
QY 709 GATTCGGCTTTACCTGAGCAGACGAGCGGAGCTTTATCTAAGCTATTTGACAGAGA 768
Db 206 AspLeuGlyTrpHisLeuSerArgGlnThrGlyGlyLeuThrArgAlaMetAspArgGly 225
QY 769 ACAAGGGATAGCTTTTGGTCCGAGTGGCTTGAATTTAATCTCTCCATCATGTTT 828
Db 226 ThrLysGlyIleSerGlnValLeuIleAsnValPheHisIleIleProIleSerPhe 245
QY 829 GAAGTATGCTTGTCAGTGGTGTGTTGTATTCAAATGCGGTGCCAGATTGCTTGTA 888

```

Db 246 GlutSerValIcylGlyLeuThrTrpGlnPheGlyAlaSerPheAlaIle 265
Qy 889 ACCCTGGACAACCTGGATACACAGACATTCACAGTGCAGTACACGGGGAAGT 948
Db 266 ThrPheSerThrMetLeuLeuTrpSerIlePheThrIleValThrAlaTrpThr 285
Qy 949 AGATTAGAAATGAGATGACCAAGACAGATTAATGATGACAGTAATGCTGTATGACCA 1008
Db 286 HisPheArgAlaAspAlaAsnLysAlaAspAsnLysAlaAlaSerValAlaLeuAspSer 305
Qy 1009 CTGTGAATTATGAAACCTGTGAAGTATTTTAATGAAAGATATGAAACAGAGATAT 1068
Db 306 LeuIleAsnPheGluAlaValLysTrpPheAsnAsnGluLysTyrLeuAlaAspLysTrp 325
Qy 1069 GATGATTTTGAACAGCTGATGAGCTGCTTATGAAAGACCTTACTCTGCTATG 1128
Db 326 AsnGlySerLeuMetAsnLysTrpArgAspSerGlnIleLysValSerGlnSerLeuAlaPhe 345
Qy 1129 CTGAACCTTGGTCAAGTCTATTTTCAGTCCGTTTACAGCTATATGCTGCTCGCC 1188
Db 346 LeuAsnSerGlyGlnAsnLeuIlePheThrAlaLeuThrAlaMetCysTrpGly 365
Qy 1189 AGTCAGGAAATGTGCAAGTACCTTACTGTGAGATCTATGATGATGATGAGACTG 1248
Db 366 CysThrGlyValIleGlyGlyAsnLeuThrValGlyAspLeuValLeuIleAsnGlnLeu 385
Qy 1249 CTTTTCAGCTTTCATTAACCTGTAACCTTTCGGAACTGTATATAGAGACTGACAA 1308
Db 386 ValPheGlnLeuSerValProLeuAsnPheLeuGlySerValTrpArgAspLeuLysGln 405
Qy 1309 GCATCATGATATGATGAACCTGTGTTACTTACTTACTCAAGTATGACCAACCAATTAAGAC 1368
Db 406 SerLeuIleAspMetGluThrLeuPheLysLeuArgLysAsnGluValLysIleLysAsn 425
Qy 1369 AAAGTATGATGATCTCCCTTCAGATCAACCCAGACAGACT-----ACCGTGGCTTT 1422
Db 426 -----AlaGluArgProLeuMetLeu---ProGluAsnValProLysAspIleThrPhe 442
Qy 1423 GATATGTGATTTTGAATACATTGAGGGCCAGAAAGTCTTATAGTGAATACCTTTGAA 1482
Db 443 GluAsnValIleThrPheGlyTrpIleAspProAspArgLysIleLeuLysAsnAlaSerPheThr 462
Qy 1483 GTCCCTGAGGAGAAAGAAAGTGGCATTTGAGAGGATGGTCAAGGAGGAAAGCAATA 1542
Db 463 IleProAlaGlyTrpLysThrAlaIleValGlySerSerGlySerGlyLysSerThrIle 482
Qy 1543 GTAGAGCTATTAATTTGCTTCTATGAGCCTCAAAAGGATGATTTATCTTGCTGCA 1602
Db 483 LeuLysLeuValPheArgPheTrpAspProGlnSerGlyArgIleLeuIleAsnGlyArg 502
Qy 1603 AATATACAGATGTGAGCCTGGAAGCCTTCGAGGGCAGTGGAGTGTACTCAGAT 1662
Db 503 AspIleLysGluTrpAspIleAspAlaLeuArgLysValIleGlyValValProGlnAsp 522
Qy 1663 GGTGCTCTCTTCATTAATCTTATTAACAACCTCTTATGAGAACTCAAGCTTCA 1722
Db 523 ThrProLeuPheAsnAspThrIleTrpGluAsnValLysPheGlyArgIleAspAlaThr 542
Qy 1723 CTTGAGAAAGTATGAGCTGGAATTAAGCTGCAATTCATGATGCAATTTCTGCAATG 1782
Db 543 AspGluGluValIleThrValAlaGluLysAlaGlnLeuAlaProLeuIleLysLysLeu 562
Qy 1783 CCACATGATATGACCAACCAAGTGGGAGAGAGACTCAGCTTTCAGAGAGAGAAAG 1842
Db 563 ProGlnGlyPheAspThrIleValGlyGluArgGlyLeuMetIleSerGlyGlyLys 582
Qy 1843 CAAAGATGACATTCAGACAGCACTTTTGAAGAGACCCCACTCATCTCATATGATGA 1902
Db 583 GlnArgLeuAlaIleAlaArgValLeuLeuLysAsnAlaArgIleMetCysPheAspGln 602
Qy 1903 GCTACTCATCGTTAGATTGATTACTGAAGACATTTCTTGCTGCAATGAAGAT-- 1959

```

```

Db 603 AlaThrSerAlaLeuAspThrHisThrGluGlnAlaLeuLeuArgThrIleArgAspAsn 622
Qy 1960 ---GTGGCAAAACACAGAACTCTATTCTTCTGACACAGATTTGCAACAGTGGTGA 2016
Db 623 PheThrSerGlySerArgThrSerValTrpIleAlaHisArgLeuAlaGlnThrIleAlaAsp 642
Qy 2017 GCAGATGAATCATTTGTTGGATCAGGGTAAAGGTACCCGAACTGTATCCACCATGCT 2076
Db 643 AlaAspLysIleIleValLeuAspAsnGlyArgValArgGluGluGlyLysHisLeuGlu 662
Qy 2077 TTGCTGTATACCTCATAGTATCTATTCAGAAATGTGGCATACACAGACAGCCGCTG 2136
Db 663 LeuLeuAlaMetProGlySerLeuTrpArgGluLeuTrpThrIleGln----- 678
Qy 2137 CAGAACATGATTAACCCCAATGGAGCAAGAAAGAAATATATATCAAAAGACAGGAA 2196
Db 679 -----GluAspLeuAspHisLeuGluAsn 686
Qy 2197 AGAAAGAACTACAGAAATAT 2220
Db 687 GluLeuLysAspGlnGlnGluLeu 694

RESULT 6
US-09-195-391-4
: Sequence 4, Application US/09195391
: Patent No. 6080842
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purni
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
: TITLE OF INVENTION: TRANSPORT PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/195,391
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/895,522
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0336 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 694 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 575393
: US-09-195-391-4

Alignment Scores: 4.17e-147 Length: 694
Pred. No.:

```

Score: 1516.50 Matches: 304
 Percent Similarity: 67.83% Conservative: 122
 Best Local Similarity: 48.41% Mismatches: 169
 Query Match: 35.65% Indels: 33
 DB: 3 Gaps: 7

AF133659 (1-2345) x US-09-195-391-4 (1-694)

QY 355 AAAATCATAAAGCAATGCTTTCTTATGTGTGCCAAAGACAGCCAGATCTACGAGCT 414
 Db 94 LyslelleuysaspheupheAIGTyrIletrProlysglyAenAmlValArgIle 113
 QY 415 AGAGTTGCCATTTGCTGGGATTTTGGGGTGGCAAGCCATGAATATGCTGCC 474
 Db 114 ArgValleullealaleuGlyleuLeulleSerAlaIlylleuAenValGlnValPro 133
 QY 475 TTGATGTTAAATATGCTGTAGACAGCTCAACAGATGTGGGAAACATGCTGAACCTG 534
 Db 134 PhephephepysglnthrIleApsSer-----MetAmlle 145
 QY 535 AGTATGACCAAAATACAGTTGCAACCATGACACA-----GCAGTTGATGAGCTAT 588
 Db 146 AlatrphaspaprophthrValAlaLeuProAlaAlaleGlyleuThrIleleuGlyTyr 165
 QY 589 GGTGTATCAAGAGCTGAGCTGCTTTTAAAGAGTGGAAATGCAATATTTGGCAAG 648
 Db 166 GlyValAlaIarPheglySerValleuPheGlyGluLeuAraGlnAlaValPheAlaIlys 185
 QY 649 GTAGCCCAAGATTCATCCCAAGAAATAGCCAAATATCTTTCCCTCCATTCACAACTG 708
 Db 186 ValAlaGlnAenAlaIleAraGlnValSerleuGlnThrPheGlnIleleuMetLysLeu 205
 QY 709 GATCTGGGTTTTCACCTGACAGACAGACGAGCTTATCTAAAGCTATTGACAGAGA 768
 Db 206 AspleucllytrphlsleuserAraGlnThrGlyleuThrAraGlnAmeAaspAraGly 225
 QY 769 ACAAGGGATACAGTTTGTCTGAGTGTGGTATTAATCTTCTCCATCANGTTT 828
 Db 226 ThrlysglyIleSerGlnValleuThrAlaMetValPhehIleleProIleSerPhe 245
 QY 829 GAAGTATGCTGTCACTGAGTGTGTTTGTATTAACAATGCCGTCGCCAGTTTGTGTA 888
 Db 246 GluIleSerValValCysGlyIleLeuThrIleGlnPheGlyAlaSerPheAlaAlaIle 265
 QY 889 ACCCTTGAACACTTGTATACATACAGACTTACAGTTCAGTACAGCCGTTGAGAACT 948
 Db 266 ThrPheSerThrMetleuLeuTyrSerIlePheThrIleGlyThrAlaIlePheGlnThr 285
 QY 949 AGATTGAATAGAAATGAACAAGCATATGATGATGATGATGATGATGATGATGAT 1008
 Db 286 HisPheAraGlnAraPheAlaAenLysAlaApsAmlAlaIleSerValAlaLeuApsSer 305
 QY 1009 CTCTGGAATATGAACCTGGAAGTATTTATTAATTAATGAAGATATGAAGACAGATAT 1068
 Db 306 LeuIleAenPheGlnAlaValLysTyrPheAraAmlLysTyrLeuAlaApsLysTyr 325
 QY 1069 GATGATTTTGAAGACGTTGAGACTGCTTCAATTGAAAAGTCACTTCACTGAGCTATG 1128
 Db 326 AsnGlySerleuMetAenTyrAraGspSerGlnIleLysValSerGlnSerleuAlaPhe 345
 QY 1129 CTGAATCTTGTCAAACTGCTATTTTCACTGCTGCTTAAACAGTATATGCTGCTGCC 1188
 Db 346 LeuAraSerGlyGlnAenLeuIlePheThrThrAlaLeuThrAlaMetLysTyrMetGly 365
 QY 1189 AGTCAGGAATTTGTGGAGTACCTTACTGTTGGAGATCTAGTATATGGAATGAGCTG 1248
 Db 366 CysThrGlyValIleGlyGlnAenLeuThrValGlyAraPheValleuIleAenGlnLeu 385
 QY 1249 CTTTTCAGTTTCATACCTGAGCTTCTGGAACCTATATATGAGACATAGACAA 1308
 Db 386 ValPheGlnleuSerValProleuAraPheleGlySerValTyrAraPheLysGln 405
 QY 1309 GCATCATATGATGAACACCTTGTCTTACTCAAGGTAGACCCCAATTAAGAC 1368

Db 406 SerleulleaspMetGluThrleuPheLysleuAraGlyAraGlnValLysIleLysAra 425
 QY 1369 AAGATGATGCATCTCCCTTCAGATCAACACAGACAGCT-----ACCGGGCCTTT 1422
 Db 426 -----AlaGluAraProleuMetLeu---ProGlnAenValProTyrAraPheIleThrPhe 442
 QY 1423 GATAATGTCATTTGAATATCATTTGAGGGCCAGAAAGTCCCTTATGTAATATCTTTGAA 1482
 Db 443 GluAenValThrPheGlyTyrHisProAraPheGlyAlaIleuLysAraAmlAserPheThr 462
 QY 1483 GTCCCTGCAGGAAAGAAAGTGGCATTGTAAGAGGTATGGGTGTCAGGAAAGACACAATA 1542
 Db 463 IleProAlaGlyTyrPysThrAlaIleValGlySerSerGlySerGlyLysSerThrIle 482
 QY 1543 GTAGGCTATTAATTTGCTTCTATGAGCCTCAAAAGGTATGACATTAATCTGCTGTCGA 1602
 Db 483 LeuLysleuValPheAraPheTyrAraPheProGlnSerGlyAraGlyIleleuIleAenGlyAra 502
 QY 1603 AATATACAGATGTGAGCCTGGAAGCCTTGGAGGGCAGTGGAGTGTACTACAGAT 1662
 Db 503 AsplleLysGluTyrAraPheAlaAlaLeuAraGlyValIleGlyValValProGlnAraP 522
 QY 1663 GCTGCTCTTCCATTAATCTATTTATTAACAACCTTTATATGAAACATCACTGCTTCA 1722
 Db 523 ThrProleuPheAraAraPheThrIleTyrGlnAenValLysPheGlyAraIleAraIleThr 542
 QY 1723 CCGTAAGAGTGTATGACAGCAAAATAGCTGACCTTCATGATGCAATCTTCCGAATG 1782
 Db 543 AspGlnGluValIleThrValValGluLysAlaGlnLeuAlaPheleuIleLysLysLeu 562
 QY 1783 CCATATGATATGACACCAAGATGAGGAAAGAGAGCTCAAGCTTCAGAGAGAGAAAG 1842
 Db 563 ProGlnGlyPheAraPheThrIleValGlyGluAraGlyLeuMetIleSerGlyGlyLys 582
 QY 1843 CAAGAGTACCAATTCAGAGAGCATTGTAAGAGACCCCACTCACTACTATGATGA 1902
 Db 583 GlnAraLeuAlaIleAlaAraGlyValleuLysAraAlaAraGlyLeuMetPheAraGln 602
 QY 1903 GCTACTTCACTGTTAATTCGATTAATGATTAAGAGACATTTCTGTGTCCTCAAGAAAT 1959
 Db 603 AlaThrSerAlaLeuAraPheThrIleGlnAlaLeuAraGlnThrIleAraAraPhe 622
 QY 1960 ---GTGTCAAACACAGACTTCTATTTTCACTGACACAGATGTCACAGAGTGTGAT 2016
 Db 623 PheThrSerGlySerAraGlnThrSerValTyrIleAlaHisAraGlnAraGlnThrIleAlaAraP 642
 QY 2017 GCAGATGAATATCATTTGTCTGATCAGGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2076
 Db 643 AlaAraPheIleIleValleuAraPheAraGlnAraGlnAraGlnAraGlnAraGlnAraGln 662
 QY 2077 TTGCTGCTCAACCTCATATGATCTATTTCAAGAAATGTGGCATACAGAGACAGCCGTG 2136
 Db 663 LeuLeuAlaMetProLysSerleuTyrAraGlnLeuThrIleGln----- 678
 QY 2137 CAGAACCATGATTAACCCCAATGAGGAAAGAAAGAAATATATCAAAAGAGAGAA 2196
 Db 679 -----GluAraPheAraPheIleGlnAraPheIleGlnAraPheIleGlnAraPhe 686
 QY 2197 AGAAAGAACTACAGAGAAATTT 2220
 Db 687 GluLeuLysAraPheGlnGlnLeu 694

RESULT 7
 US-09-212-247C-4
 ; Sequence 4, Application US/09212247C
 ; Patent No. 6391603
 ; GENERAL INFORMATION:
 ; APPLICANT: POMPEYUS, Markus; SUEHLBERGER, Harald; JOSEFFKEN, Hans
 ; WOLFGANG, DOVAL, Jose Luis Raveletta; JIMENEZ, Alberto;
 ; and GARCIA, Maria Angeles Santos
 ; TITLE OF INVENTION: Genes of purine biosynthesis from AsbAya Gossypii
 ; and the use thereof in microbial riboflavin

```

;
;      synthesis
;      NUMBER OF SEQUENCES: 21
;      CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Keil & Weinkauff
;      STREET: 1101 Connecticut Avenue
;      CITY: Washington
;      STATE: D.C.
;      COUNTRY: USA
;      ZIP: 20036
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER: IBM AT-compatible, Pentium processor
;      OPERATING SYSTEM: Windows 95
;      SOFTWARE: Wordperfect version 6.1
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/212,247C
;      FILING DATE: 16-Dec-1998
;      CLASSIFICATION: <unknown>
;
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 475 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
;      US-09-212-247C-4
;
;
;      Alignment Scores:
;      Pred. No.:      8,45e-115      Length:      475
;      Score:          1201.50         Matches:      234
;      Percent Similarity: 71.74%      Conservative: 96
;      Best Local Similarity: 50.87%    Mismatches:  127
;      Query Match:      28.24%        Indels:       3
;      DB:                4            Gaps:         2
;
;      AF133659 (1-2345) x US-09-212-247C-4 (1-475)
;
;      QY 757 ATTGACAGAGCAAGGGGATGATGCTTGTCTGAGTGGTTGATTTAATCTTCT 816
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 1 MetaphaeglyCyalyseglylleserlyrvalleuserAlaMetValPhehisIleIle 20
;
;      QY 817 CCCATCATGTTTGAAGTGATGCTTGTCAAGTGCTGTTTGTATTACAATGCGGTCCGAG 876
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 21 ProlethrPhegluIleaserMetValCysglylIleleuthrTyrglnPheglIylaser 40
;
;      QY 877 TTTCGTTGGTACCCCTTGGAACACTTGATACATACAGCATTCACAGTTTCAGTCA 936
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 41 PheAlaIleIlethrPheSerThrMetLeuLeuTySerIlePheThrPheArgThrThr 60
;
;      QY 937 CGGTGAGAGATGATTTGAATGAATGAACAAGAGATATGATGATGATGATGCT 996
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 61 AlatrPheglnThrArgPheArgPheArgPheAlaAsnlyAlaAspAsnlyAlaAlaSerVal 80
;
;      QY 997 GCTATAGACTGCTGCTGAATTAATGAACCTGGAAGTATTTTAATGAATGAAGATATGA 1056
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 81 AlaLeuAspSerLeuIleAsnPhegluAlaVallyTyPheAsnAsnIulustyIleu 100
;
;      QY 1057 GCACAGAGATATGATGATTTTGAAGAGCTGTGAGACTGCTTCATGAAAAGTACCTCT 1116
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 101 AlaAspLyTyHisIleThrSerLeuMetLyTyArgAspSerGlnIleIlySerValSerGln 120
;
;      QY 1117 ACTCGGCTATGCTGATGCTGCTCAAGTGTATTTTTCAGTGGTGGTTAAGCTATTA 1176
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 121 SerLeuAlaPheLeuAsnThrIleGlnAsnLeuIlePheThrAlaLeuThrAlaIleC 140
;
;      QY 1177 ATGTGCTCGCCAGTACGGAATTTGGCAGGTACCTTACTGTGTGAGATCTAGTAATG 1236
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 141 MetTyrrMetAlaCysAsnIylValMetGlnIlySerLeuThrValGlyAspIleuValIleu 160
;
;      QY 1237 GTGAATGACCTGCTTTTTCAGCTTTTCATTAACCCCTGGAACCTTTTGGGAACTGTATATGA 1296
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 161 IleAsnGlnLeuValPheGlnIleuSerValProLeuAsnPheLeuGlySerValTyArg 180

```

```

;
;      QY 1297 GAGACTAGACAGACACTCATATGATATGAACACTTGTTTACTTACTCAAGTAGACACC 1356
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 181 AspleuIysGlnSerLeuIleAspMetGluSerLeuPheIlySerGlnIlyAsnGlnVal 200
;
;      QY 1357 CAATTTAAAGCAAAAGTATGGCATCTCCCTTCAAGTATCAACACAGACATCAACCTG 1416
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 201 ThrIleIysAsnSerProAsnAlaGlnAsnLeuProIle--HisIysProLeuAspIle 219
;
;      QY 1417 GCCTTGATATATGTCATTTTGAATCATTCAGGCGCCAGCAAGCTCTTAGTGAATATCC 1476
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 220 ArgPheGlnAsnValThrPheGlyTyAspProGluArgArgIleLeuAsnValSer 239
;
;      QY 1477 TTTCAGTCCCTGACGAGAAAGAAAGTGGCCATTGTAGACGACTGCTGACGAGAAAGC 1536
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 240 PheThrIleProAlaGlyMetLeuTyThrAlaIleValGlyProSerGlySerGlyIlySer 259
;
;      QY 1537 ACAATATGAGAGCTATTTATTTTCGCTTCTATGAGCCTCAAAAAGGTATACATTATCTTCT 1596
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 260 ThrIleLeuIysLeuValPheArgPheTyrgIuProGlnGlnIlyArgIleLeuValGly 279
;
;      QY 1597 GGTCAAAATATACAGATGTGAGCGCTGAAAGCCTTGAGAGGCGAGTGGAGTGTACT 1656
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 280 GlyThrAspIleArgAspLeuAspLeuSerLeuArgIlyAlaIleGlyValValPro 299
;
;      QY 1657 CAGATGCTGTCTCTTCATATATATCTATTATTAACAACCTCTTATATGAGAAACATGAGT 1716
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 300 GlnAspThrProLeuPheAsnAspThrIleTyrGlnAsnValIlyPheGlyAsnIleSer 319
;
;      QY 1717 GCTTCACCTGAGGAAGTATGACAGTGGCAAAATTAAGCTGAGCTTCATGATGCAATCTT 1776
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 320 SerSerAspAspGluIleLeuArgAlaIleGluIlyAlaGlnLeuThrIlyLeuLeuGln 339
;
;      QY 1777 CGAATGCCACATGATATGACACCCAGTAGGAGGAGAGACTCAAGCTTTCAGAGAGA 1836
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 340 AsnLeuProIysGlyAlaSerThrValValGlyIuArgIlyLeuMetIleSerGly 359
;
;      QY 1837 GAAAGCAAAAGATGAGCAATTTGCAAGGCCATTTTGAAGAGCCCGCATCATCTAT 1896
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 360 GluIysGlnArgLeuAlaIleAlaIleArgValLeuLeuIlyAspAlaProLeuMetPhePhe 379
;
;      QY 1897 GATGAGCTACTTATGTTTGATTCATTTCTGAAGAGACTATTTTGGTCCATGAG 1956
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 380 AspGluIleThrSerAlaLeuAspThrHisThrGlnIleAlaLeuHisIleGln 399
;
;      QY 1957 GAT-----GTGTCAAACACAGAACTTTATTTTCAATTCAGACAGATGTCAACAGTG 2010
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 400 GlnAsnPheSerSerAsnSerLyThrSerValIlyValAlaHisArgLeuArgThrIle 419
;
;      QY 2011 GTTATGACAGATGAATCATTTGCTTGATCAAGGTAAGGTAAGCCGAACGTGTACCCAC 2070
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 420 AlaAspAlaAspLyIleIleValLeuGlnGlnIlySerValArgGlnGlnIlyThrHis 439
;
;      QY 2071 CATGTTGCTTGTCTTAACCCCTCATAGTATCTATTCAGAAATGTGGCATACACAGACG 2130
;      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
;      DB 440 SerSerLeuLeuAlaSerGlnIlySerLeuTyArgIlyLeuTyPaspIleGlnIlyAsn 459
;
;      RESULT 8
;      US-09-134-001C-4730
;      Sequence 4730, Application US/09134001C
;      Patent No. 6380370
;      GENERAL INFORMATION:
;      APPLICANT: Lynn Doucette-Stamm et al
;      TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;      TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;      FILE REFERENCE: GTC-007
;      CURRENT APPLICATION NUMBER: US/09/134,001C
;      PRIOR FILING DATE: 1998-08-13
;      PRIOR APPLICATION NUMBER: US 60/064,964
;      PRIOR FILING DATE: 1997-11-08
;      PRIOR APPLICATION NUMBER: US 60/055,779
;      PRIOR FILING DATE: 1997-08-14
;      NUMBER OF SEQ ID NOS: 5674
;      SEQ ID NO 4730

```



```

; LENGTH: 582
; TYPE: PRF
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4730

Alignment Scores:
Pred. No.: 1.07e-62 Length: 582
Score: 695.50 Matches: 186
Percent Similarity: 51.41% Conservative: 123
Best Local Similarity: 30.95% Mismatches: 261
Query Match: 16.35% Indels: 31
DB: Gaps: 12

AF133659 (1-2345) x US-09-134-001C-4730 (1-582)

QY 352 CGGAAATCATGAAAGAAATGCTTTCTTATGTTGCGCCAAAGACAGGCGATCTACGA 411
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ArgArgnctleleuysarglyrleuylphevalyspro-----TyrArg 17

QY 412 GCTAGAGTGGCCATTTGCGTGGGATTTTGGGTGCGCAAG--GCCATGAATATTGTG 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 18 TyrArgilelealeatrhile---llevalgilylellelelyspheglylleprometleu 36

QY 469 GTTCCCTTACGTTAAATATGCTGTAGACAGCTCAACGATGCGGAAACATGCTG 528
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 37 lleproleuileuileuylsryalalleasp-----Glyvalille 49

QY 529 AACCTGAGTATGACCAATATACAGTTCACATGCGAACAGAGCTTCGATTGGC--- 585
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 50 Asnasmhisserleuthrphenglnglulyspheserhisleuglyvalalleglylle 69

QY 586 -----TATGTTATACAGAGCTGAGCTGCTTTTAAAGAGTTGCAAT 633
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 AlaleupheillepheleullevalArgproproilleugluphe-----lleArgln 86

QY 634 GCATATTTGGCAGAGTACCAGAAATTCGAATCCGAAAGATGCGCAAAATGCTTCTC 693
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 Tyreulnlaeintrprhiserasnlyseilleuutyrspllearglysglnleutyasn 106

QY 694 CATCTCAACAACCTGATCTGGGTTTTCACCTGAGCAGACAGACAGGGA---GCTTATCT 750
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 107 HisleuqlnalleuservalArgpheTyralaasnnglnvalglglnvalilleSer 126

QY 751 AAGGCTATGACAGAGAACAGGGGATACAGTTTGTCTGATGCTTGTGTAATTTAT 810
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 Argvalilleasnvalglnglnthrlyasphelleleuthrlyleuemetasnille 146

QY 811 CTCTTCCCATCATGTTTGAAGTATGCTGTCAGTGTGTTTGTATTCAAAATGCGGT 870
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 147 TrpleuaspCysillethrillellelleallealeuserillemetPhepheleuaspvallys 166

QY 871 GCCAGATTTGCTTGTGTAACCTTGAAACACTTGTAATACAGAGA---TTCCAGGTT 927
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 167 LeuthrPhealleallepheillePheproPheTyrlleuthrvalTyPhephephe 186

QY 928 GCATCTACACGCTGAGAGACTAGATTAGATAGAAATGAACAAGACAGATATGATGCA 987
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 GlyArgleuArglyleuthrArgvalArg-----serglnalleuallagluval 203

QY 988 GGTATGCTGCTATGACTACTGCTGAATTTGAAACTGTAAGATTTTAAATATGAA 1047
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 GlnlglyPheleuhsigluargvalglnglymetservalilleuyserserPheallelelu 223

QY 1048 AGATATGAGACAGAGATATGATGATTTTGAAGACGATGAGATGCTTCACTTGA 1107
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 224 AspsnnglnlallayasnphesasnhisasnlyasnphesleuglnarglalaPhegln 243

QY 1108 AGTACCTTACTGCTGCTATGCTGAATTTGTGTCAAAGTCTATTTTCAAGTGTGTTA 1167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 HisthrArgtrpansnialatyserPheallealleasnthrvalThrAspleuuglyPro 263

QY 1168 ACAGCTAAATGGTCTGCCGACAGGGAATTTGGCAGGATCCCTTACTGTTGGAGAT 1227
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 264 llellevalilleglyvalglyserTyPheuallallethrGlyserillethrvalglyThr 283

QY 1228 CTAGTAATGCTGATGAGTACTGCTTTTTCAGCTTTTCACTTATACCCCTGAACCTTCTGGAGACT 1287
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 leuallalephevalglyTyPheleuglnglnleuPheoglyProleuamrAgaleuvalser 303

QY 1288 GTATATGAGAGACTGACAGACATCATATGATATGAAACCTTGTTTTACTTACTCAAG 1347
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 SerPheThrThrleuthrGlninserPhealleaserMetaspArgvalPheglnleuMetasp 323

QY 1348 GTAGACACCCAAATTTAAGCAAGATGATGCTTCCTCCCTTCAAGTACACACACAGACA 1407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 GluAspTyrsplleleuysasnlglylleglyalaglnProilleysilleser-----Lys 341

QY 1408 GCTACCGTGCCTTTGATATATGATGCAATTTGATATACATTGAGGCGCAGAA---GTCCTT 1464
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 342 GlyGlnlleaspleuysnhsvalserPheleTyTrasnlglnsnlglysgluvalleu 361

QY 1465 AGTGAATATCTTTGAAAGTCCCTGACAGAAAGAAAGTGCCATTTGAGAGGTAGTGGG 1524
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 HisAspilleasnleuthrilleasnlysglygluThrValallephevalglymetsergly 381

QY 1525 TCAGGAAAGACCATATAGAGGCTATATTTGCTTCTATAGAGCTCAAAAAGGTAGC 1584
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 382 GlyglylyserthrleuileasnleulleProArgPheTyrsvalThrnglnglylu 401

QY 1585 ATTATCTTCTGCTGCTCAAAATATACAGATGTAGCGCTGGAAGCCTTGAGAGGGCAGTG 1644
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 402 lleleuileasphishisasnvallyspasphleuthrglyserleuamragnlnille 421

QY 1645 GAGGTGATCCTGAGAGTGTGCTCTTCCATATATCTATTTATTAACAACCTTATAT 1704
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 GlyleuvalGlnlnasphenilleleupheserAspThrVallysglnsnilleleu 441

QY 1705 GGAATATCATGCTTCACTGAGAGAGTATATGACGTGCAAAATTAAGTGCATTTCA 1764
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 442 GlyArgProaspallethrspaspgluvalvalgluallaleuysmetalaaenlallhis 461

QY 1765 GATGCAATCTTCTGATGCGACATGATATGACACCCAGTAGGGAACGAGAGCTCAAG 1824
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 462 AspPheilleaserasnleuProasnlglyTyrsPthrgluvalglygluarglyvallys 481

QY 1825 CTTCAGAGAGGAAAGAAAGAGTAGACATTTGCAAGCCATTTTGAAGACCCCA 1884
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 leuserglyglylnlysglnargleuserillealargillepheleuasnAnproPro 501

QY 1885 GTATATCTATATGAGTACTTATGATTGATTCGATTACTGAAGACTATTTCTT 1944
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 502 Valleuilleuaspgluallathrseralleuaspbleuclusergluallalleleln 521

QY 1945 GGTGCCATGAGAGATGTCGCAACACAGAACTTCTATTTTTCATGACACAGATGTCA 2004
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 522 Gluallaleuaspvalleuserlyspaspglnthrleuilevalalleahisargleuser 541

QY 2005 ACAGTGTGATGACAGATGAATCATTTGTGATGACAGGTAAGGTAGCCGATGCTG 2064
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 ThrillethrHisallaspArgllevalValmetGlnasnlglyArgllevalgluthrgly 561

QY 2065 ACCCACCATGCTTGTCTGCTAACTCTCATATGATATTTACGAATGTGGCATACAG 2124
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 ThrHisGlnlglnleuile---AsnlyAsrglyalatygluhisleutyserillegln 580

QY 2125 AGC 2127

Db 581 Asn 581

RESULT 9
US-08-772-270A-4
; Sequence 4, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul

```



```
Db 630 AspglnalThrSerAlaLeuAspArgLysSerGlnHisIleMetArgAsnMetHis 649
Qy 1957 GATGTGCAACACAGACTTCTATTTCATGACACAGACTGTCAAGCTGGTAT 2016
Db 650 GlnIleCysArgLysGlnThrValIleIleIleAlaHisArgLysSerThrValLysAsn 669
Qy 2017 GCAGATGAATATCATGTTGCTTGATCAGGGTAGAGTACGCCGAACGTGTATCCACCATGCT 2076
Db 670 AlaAspArgIleIleValMetGlnLysGlnIleValGlnGlnGlyLysHisLysGln 689
Qy 2077 TTGCTTGCTAACCTCATAGTATCTATTCAGAAATGTGGCATACAGAC 2127
Db 690 LeuLeuAlaAspProAsnGlyLeuThrHisIleTyrLeuHisGlnLeuGlnSer 706

RESULT 10
US-09-134-001C-5611
; Sequence 5611, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5611
; LENGTH: 580
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5611

Alignment Scores:
Pred. No.: 9,86e-57 Length: 580
Score: 637.50 Matches: 174
Percent Similarity: 50.52% Conservative: 118
Best Local Similarity: 30.10% Mismatches: 239
Query Match: 14.99% Indels: 47
DB: 4 Gaps: 12

AF133659 (1-2345) x US-09-134-001C-5611 (1-580)
Qy 406 CTACGACCTAGAGTGGCATTGCTGGAGATTTTGGGTGGCAAGGCCATGATATT 465
Db 27 TLeAlaAlaValIleIleIleThrSerLeuGlySerLeuSergly-----Leu 41
Qy 466 GTGGTTCCTCC-----TTTCATGTTAAATATGCTGTAGACAGCCCTAAC--- 507
Db 42 LeuValProLeuPheThrGlyArgLeuValAspLysPheSerValSerSerIleAsnTyr 61
Qy 508 -----CAGATGTCGGGAACATGCTGAACCTGAGTATGACCAAAATACAGTTGCA 558
Db 62 GlyMetIleAlaIlePheGlySerIlePhe----- 71
Qy 559 ACCATGCGAACAGCAGATTGCTGATTGGCTATGCTGTATCAGAGCTGAGCTGCTTTT 618
Db 72 ---LeuValAsnAlaLeuLeuSerglyIleGlyLeu----- 82
Qy 619 ACGAAGATTGAATGAGATATTGGCAGAGGTACCCAGAAATTCAATCCAGAAATAGCC 678
Db 83 -----TyrLeuLeuSerLysIleIleGlyLysIleIleTyrAlaIleArg 97
Qy 679 AAAAATCTTTCTCATCTTCAACACCTGGATCTGGATTTCACCTGAGACAGACAG 738
Db 98 SerLeuLeuThrGlnHisIleIleGlnLeuLysMetProPheAspLysPheAsnGlnSer 117
Qy 739 GGAAGCTTATCTAAAGCTATTGACAGAGAAACAAGGGGTATCACTTTTGTCTGAGTCT 798
```

```
Db 118 GlyGlnLeuMetSerArgLeuThrAspAspThrLysValIleAsnGlnPheIleSerGln 137
Qy 799 TTGGATTTTATCTTCTCCCATCATGTTTGAAGTACGCTTGTCAGTGGTGTAT 858
Db 138 LysLeuProAsnLeuLeuProSerValLeu---ThrLeuIleGlySerLeuValMetLeu 156
Qy 859 TACAATGCGGTGCCAGTTTGGTTGTAACCTTGGAACTTGTATACATACACAGCA 918
Db 157 PheIleMetAspThrLysLeuThrLeuThrPheIleThrIleProValPhe----- 174
Qy 919 TTACAGCTTGACATCAGCAGGTGAGACAGAACTTATAGATTAAGAAATGAAACAAAGCAT 978
Db 175 IleLeuIleIleValProLeuGlyArgValMetGlnLysIleSerThrAsnThrGlnSer 194
Qy 979 AATGATCAGCTAATGCTGTATA-----GACTCAGCTGAAATTATGAACGTGAG 1032
Db 195 GlnIleAlaAsnPheSerGlyLeuLeuGlyArgValLeuThrGlnMetArgLeuValLys 214
Qy 1033 TATTTTAATGAATGAAGATATGAAGACACAGATATGATGATTTTGAACGTATGAG 1092
Db 215 ValSerAsnThrGlnThrGlnLeuGlnLeuAspAsnAlaHisThrAsnLeuLysIleTyr 234
Qy 1093 ACTGCTTCATTGAAAAGTACTTACTCTGCTATGCTGAACCTTGTCAAAGTCTATT 1152
Db 235 ArgLeuGlyLeuLysGlnAlaLysIleSerAlaValValGlnProIleSerglyValVal 254
Qy 1153 TTGAGTGTGCGTTTAAACAGCTATATATGCTGCTGCCAGTACAGGAATTGTGGCAGTACC 1212
Db 255 MetLeuLeuThrIleAlaIleIleLeuGlyPheGlyAlaLeuGlnIleGlyThrGlyAla 274
Qy 1213 CTTACTGTTGGAGATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db 275 IleThrProGlyThrLeuIleAlaMetIlePheTyrValIleGlnLeuSerMetProLeu 294
Qy 1273 AACTTCTGGAAGCTGATATATAGAGACTAGACAAAGCAAGCTCATAGATATGAACACTTG 1332
Db 295 IleAsnLeuSerThrLeuValIleThrAspTyrLysLysAlaValIleLysSerArgIle 314
Qy 1333 TTACTCTACTCAAGGTAGACACCCAAATTAAGCAAAAGTATGATGATGATGATGATGATGAT 1392
Db 315 TyrGlnIleMetGlnGlnProIleGlu---ProThrGlnAlaLeuSerglySerLysAsp 333
Qy 1393 ATCACAACACAGACAGCTACCGTGGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1452
Db 334 ValThrIleIleAspGlyGlnLeuValPheGlnHisValAspPheLysTyr---AspVal 352
Qy 1453 CAGAAAGTCTTAGTGAATATGCTTGAAGTCCCTGCGAGAAAGAAAGTGGCCATTGA 1512
Db 353 LysLysIleLeuGlnLysAspValSerPheSerIleProGlnGlyGlnValSerAlaPheVal 372
Qy 1513 GGAAGTAGTGGGTCAAGGAAACCAATATGATGAGGCTATATTGGCTCTATGAGCCT 1572
Db 373 GlyProSerIleSerglyLysSerThrIlePheAsnLeuIleGlnThrGlyMetTyrAspIle 392
Qy 1573 CAAAAGGTAGCATTTATCTTGTGCTGTCAAATATATCAAGATGAGCCTGGAAGCCTT 1632
Db 393 GlnArgGlyAspIleLysTyrGlnAsnGlnSerIlePheAspIleProLeuSerglyTyr 412
Qy 1633 CGAAGGCACTGGAGTGAATCTCAGATGCTGTCCTCTTCATATATCTATTTATAC 1692
Db 413 ArgThrLysIleGlyTyrValMetGlnSerAsnSerMetMetSerglyThrIleAspAsp 432
Qy 1693 AACCTCTTATATGA---AACATCAGTGGTTCACCTGAGGAAGTATGATGAGGCAAAA 1749
Db 433 AsnIleLeuTyrGlyIleAsnArgLysValAspAspGlnLeuIleGlnTyrAlaLys 452
Qy 1750 TTAGCTGACTTCATGATGCAATCTTTCGATGCCACATGATATATGACACCAAGTAGGG 1809
Db 453 LeuAlaAsnCysHisAspPheIleMetGlnPheAspGlnGlyTyrThrAspThrMetValGly 472
Qy 1810 GAAAGAGACTCAAGCTTTCAGAGAGAAACAAAGAGTACCAATTGCAAGAGCCATT 1869
Db 473 GlnArgGlyLeuLysLeuSerglyGlyGlnArgGlnArgIleAspIleAlaArgSerPhe 492
```

QY 1870 TTGAAGACCCCGCTACTATCTATGTAAGCTTACCTGATTCGATTCGATTCAT 1929
Db 493 VallysaenProaspilieleuleuleuaspGluAlaThrAlaAsnleuaspserGluSer 512
QY 1930 GAAGAGCTATTCTTGTCGCATGAAGATGTGTCMAACACAGACTTCTATTTTCATT 1989
Db 513 GluleuLysIleGlnGluAlaLeuGluThrLeuMetGluGlyArgThrThrValIle 532
QY 1990 GCACACATTTCTCAACACTGTTGATGACATGAAATTCATTTCTTGATAGGCTAG 2049
Db 533 AlaHisArgLeuSerThrIleLysLysAlaGlyGlnIleValPheIleAspLysGlyGlu 552
QY 2050 GTAGCCGAAGCTGATCCACCATGTTTGTCTTAACCCCTCATAGATCTAT 2103
Db 553 ValThrGlyLysGlyThrHisHisGluLeuMetAlaSer--HisAspLysIle 569

RESULT 11
US-08-772-270A-12
Sequence 12, Application US/08772270A
Patent No. 6019984
GENERAL INFORMATION:
APPLICANT: MacInnes, Janet
APPLICANT: Ricciardi, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
PRODUCING SAME, AND THEIR USE AS VACCINES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-12

Alignment Scores:
Pred. No.: 1e-54 Length: 711
Score: 618.50 Matches: 165
Percent Similarity: 51.78% Conservative: 126
Best Local Similarity: 29.36% Mismatches: 242
Query Match: 14.54% Indels: 29
DB: 3 Gaps: 10

AF133659 (1-2345) x US-08-772-270A-12 (1-711)

QY 463 ATTGGTTCCTTCATGTTTAAATATGCTGTAGACAGCCCTCAACAGATGTCGGA--- 519
Db 173 LeuIleThrProLeuPhePheGlnValValMetAspLysValLeuValHisArgGlyPhe 192
QY 520 AACATGCTGAACCTGAGTATGACACCAATATACAGTTGCACATGCGACACAGCTTCG 579
Db 193 SerThrLeuAsn-----ValIleThrValAlaLeuSerValVal 205
QY 560 ATTGGCTATGTTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTCGAATGACGATA 639
Db 206 ValIlePheGluIleValLeuSerGly-----LeuArgThrTyrIle 219
QY 640 TTGGCAAGTAGCCCAAAATTCATCCAGAAATAGCCAAAATGCTTTCTCATCTT 699
Db 220 PheSerHisSerThrSerArgIleAspValGluLeuGlyAlaLysLeuPheArgHisIleu 239
QY 700 CACAACTGATCTGGGTTTTCACCTGAGCAGACAGACGGAGCTTATCTAAGCTATT 759
Db 240 LeuAlaLeuProIleSerTyrPheGluAsnArgArgValGlyAspThrValAlaArgVal 259
QY 760 GACAGAGGACAGAGGGATACGTTTGTCTGAGTGTCTTGATTTAATCTTCC 819
Db 260 ArgGluLeuAspGlnIleArgAsnPheLeuThrGlyAlaAlaLeuThrSerValLeuAsp 279
QY 820 ATCATGTTTGAAGTATGCTGTGTCAGTGTGTTTGTATTAACAAATGGCGCCAGTTT 879
Db 280 LeuLeuPheSerPheIlePhePheAlaValMetIrrPyrTyr-----SerProLysLeu 297
QY 880 GCTTGGTAACTTGGACCTTGTGATACATACACAGATTCACAGTTGCAGTACAGG 939
Db 298 ThrIleValIleLeuLeuSerLeuProCysTyrIleAlaIrrSerIlePheIleSerPro 317
QY 940 TGGAGAACTAGATTGATGAATAATGACAAACAGATAT--GATGACGATATGCT 996
Db 318 IleLeuArgArg--ArgLeuAspGluLysPheAlaArgAsnAlaAspAsnGlnSerPhe 336
QY 997 GCTATGACACTGCTGCTGAATTATGAACCTGTGAAGTATTTAAATGAAGATATGAA 1056
Db 337 LeuValGluSerValSerAlaIleAspThrIleLysAlaLeuAlaValThrProGlnMet 356
QY 1057 GCACAGAGATATGATGATTTTGAAGCGTATGAGACTGCTTATGAAAAGTACTCT 1116
Db 357 ThrAsnIleTrrAspLysGlnLeuAlaSerTyrValSerAlaAspPheArgValThr-- 375
QY 1117 ACTGCGCTATGCTGAACCTTGTGCTCAAGCTCTTTTCACTGCGGTTTAACGCTATA 1176
Db 376 -----ValLeuAlaThrIleGlyGlnGlnValGlnLeuIleGlnLysThrValMet 393
QY 1177 ATGGTG-----CTGCCAGTCAAGGAAATGTGCGAGGTACCTTACTGTGGAGAT 1227
Db 394 IleIleAsnLeuThrPheGlyAlaHisIleLeuValIleSerGlyAspLeuSerIleGlyGln 413
QY 1228 CTAGTATGTAATGACGCTGCTTTTCACTTTCATTAATCCCTGAACTTTTGGAACT 1287
Db 414 LeuIleThrPheAsnMetLeuSerGlyGlnValIleAlaProValAlaArgLeuAlaGln 433
QY 1288 GTATATAGAGACTAGACAGCACTCATATGATGAACACTGTTTACTCTACTCAAG 1347
Db 434 LeuTrpGlnAspPheGlnGlnValGlyIleSerIleThrArgLeuGlyAspValLeuAsn 453
QY 1348 GTAGACACCCAAATTAAGACAAAGTGATGCTCCCTTCAGATCACACAGACA 1407
Db 454 SerProThrGlu-----AsnTyrGlnGlyLysLeuSerLeuProGluIlePhe 469
QY 1408 GCTACCGTGGCTTTGATATGTCATTTGAATAC--ATTGAGGCCAGAAAGTCTT 1464
Db 470 GlyAspIleAlaPheLysHisIleArgPheArgTyrLysProAspAlaProIleIleLeu 489
QY 1465 AGTGAATATCTTTGAAGTCCCTGCAGAAAGAAAGTGGCCATTTGAGAGTAGTGG 1524
Db 490 AspAspValaLeuLeuSerValLysGlnGlyValIleGlyIleValGlyArgSerGly 509
QY 1525 TCAGGAAAACACAAATAGTAGGCTATTATTTGCTTATGAGCTCAAAAGGATAGC 1584

```
|||||
Db 510 SerGlyLysSerThrLeuThrLysLeuLeuGlnArgPheTyrlleProGlnAsnGlyGln 529
1585 ATTATCTTGCTGGCAATATATATACAGATGATGAGCCCTGGAAAGCCCTGGAGGGAGATG 1644
:::
Db 530 ValLeuIleArgGlyHisAspIleuAlaLeuAlaAspProAsnThrPheLeuArgGlnIle 549
1645 GAGAGTGAACCTCAAGATGCTGCTCTTCCATATATCTATTTATTTCAACCTCTTATAT 1704
550 GlyValValLeuGlnAspAsnValLeuLeuAsnArgSerIleArgAspAsnIleAlaLeu 569
Qy 1705 GGAACATCATGCTTCACCTGAGAGAGTGTATGCAAGTGGCAAAATTAGCTGGACTTAT 1764
Db 570 ThrAspProSerMetSerMetGluArgValIleTyrlAlaAlaLysLeuAlaGlyAlaHis 589
1765 GATGCAATTTCTTGCATGCCACATGATATGACACCCCAAGTAAAGGAAACGAGACTCAAG 1824
Db 590 AspPheIleSerGluLeuArgGluGlyTyrlAsnThrIleValGlyLeuGlyAlaGly 609
Qy 1825 GTTTCAGAGAGAGAAAGCAAGATGACCAATTGCAAGAGCCATTTTGGAAAGCCCCCA 1884
Db 610 LeuSerGlyGlyGlnArgGlnArgIleAlaIleAlaArgAlaLeuValAsnAsnProArg 629
Qy 1885 GTCAATCTCTGTGAAGCTACTTATCGTTAGATTGCTGAAAGAGACTATTCTT 1944
Db 630 IleLeuIlePheArgGluAlaThrSerAlaLeuAspTyrlGlnSerGlnIleIleIleMet 649
Qy 1945 GGTGCATGAGAGATGTGGTCAAAACAGAACTTATTTTCATTTGACACAGATTGCA 2004
Db 650 GlnAsnMetGlnLysIleCysHisGlyArgThrValIleIleIleAlaHisAspLeuSer 669
Qy 2005 ACAGTGTGTTGAGAGATGAATCATTTGTCTTGATCAGGTAGGTAGGTAGGTAGGTAGGT 2064
Db 670 ThrValLysAsnAlaAspArgGlyIleValMetGluLysGlyHisIleValGlnGlnGly 689
Qy 2065 ACCACCATGTTGCTTGTCTTAACCTCAATAGTATCTTATTCAGAAATGGCATACACAG 2124
Db 690 LysHisAsnGlnLeuLeuGlnAsnGlnLysGlyTyrlTyrlLeuAsnGlnLeuGln 709
Qy 2125 AGCAGC 2130
Db 710 SerAsn 711
RESULT 12
US-08-394-880B-2
; Sequence 2, Application US/08394880B
; Patent No. 5705352
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Multiple Drug Resistance Gene Of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company/Patent Division
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: US
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,880B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant G., Thomas
; REGISTRATION NUMBER: 35784
; REFERENCE/DOCKET NUMBER: X-9682
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-2459
; TELEFAX: (317) 277-1917
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 791 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-394-880B-2

Alignment Scores:
Pred. No.: 3.13e-52 Length: 791
Score: 594.50 Matches: 227
Percent Similarity: 42.84% Conservative: 129
Best Local Similarity: 27.32% Mismatches: 302
Query Match: 13.98% Indels: 173
Db: 1 Gaps: 32

AF133659 (1-2345) x US-08-394-880B-2 (1-791)
Qy 17 TGCAATCTTGCGCGCTGGCGCGCGCGCG----- 46
Db 33 CysValIleSerAlaArgHisSerAlaArgAsnGlyLeuIleArgGlnPheSerGlyCys 52
Qy 47 -----CTGCTTTCGAAAGCGCCGCGACCTCCGATTTCTGATCCGCTTTAGTCTCTG 100
Db 53 SerGlySerIleSerAsnSerGlyAsnProArgProTyrlArgSerAlaIleThrSerLeu 72
Qy 101 TTAGCGGCTCAGTCCGCGACGTGAGGC----- 127
Db 73 LeuSerAlaAsnValCysSerLysGlyValSerAlaValGlnProArgPheLeuSerThr 92
Qy 128 ---CACATCAACTCGGCGCTTGGAAACCGCTGAGCCT----- 163
Db 93 ValArgLeuPheSerThrSerGlnArgSerLeuGlnProLysSerAsnValLysSerThr 112
Qy 164 -----ACCAGATTCAGAGTCATTAAAGATATCAATGCGACAGATTGGGAA 211
Db 113 GlyGlyGlnValValArgProGlnLeuHisGlnAspGlnIleHis---GluAspIleGln 131
Qy 212 AAGGCAATTCAGACAGCTTTAGATGCTGCAAAAGCTCTCCAGATGAGCCACTGATAG 271
Db 132 LysGlyPheGluLeuSerGlnArg-AlaAlaGlnAlaAlaGlnVal-----AsnLeu 149
Qy 272 AAAAGAGCATGTGGCATGTGCATGTCAGAGAGAGAGATCCACAGACCCAAAAGAG 331
Db 149 exAlaLysLeuAlaLysAspGlyAlaIleGly-----LysLysAlaG 163
Qy 332 GGTAAAGATGTTGATCTCGGAAATCATTAAGCAATGCTTTCTTATGTGTGGCCCA 391
Db 163 LysPheLysGluIle---TrpArgLeuLeuIleAla----- 174
Qy 392 AAGACAGCCAGATCTACAGACTAGAGTTGCCATTTGCTGAGATTTTGGTGGTCA 451
Db 175 -----ArgProGlnAlaLys---LysLeuAlaLeuAlaPheLeuPheLeuValSer 192
Qy 452 AGGCCATGAATTTGCTGCTTCCTTCATGTTTAAATATGCTGTAGACAGCCTCAACAGA 511
Db 192 exGlyIleThrMetSerIleProPheSerIleGlyLysIleMetAspThrSerThrLys 212
Qy 512 TGTGCG-----GGAAACATGCTGAACTGAGTATGACCAAAATATACAGTTGCAACCA 562
Db 212 laThrThrArgGlyGlyAsnGlnLeuPheGlyLeuSerLeuProMetPheTyrlAla 232
Qy 563 TGGCAACAGCAGTTCTGATGCTATGCTGTATCAAGAGCTGAGAGCTGCTTTTATACG 622
Db 232 euAla-----GlyIleLeuThrLeuGlyAlaAla---AlaAsn 244
Qy 623 AAGTTGAAATGACATTTTGGCAAGTACCCAGAAATTAATCCGAAGATAGCCAAA 682
Db 244 yrlTyrlArgIleIleIleLeuArgIleValGlyGlnArgIleValAlaArgLeuArgSer 264
```

Qy	683	ATGCTTTCTCCATCTTCACAACTGGATCGGGTTT	-----	720
Db	264	ysleuPhehArgGlnThrPheValGlnmsplacIuPhePheAspAlaAsnArgValGlyA	-----	284
Qy	721	-----CACCTGACGACAGACGGAGACTTATTCATAGCCTTATGACAGA----	765	
Db	284	spLeuIleSerArgLeuSerSerAspHrIleIleValGlyysSerIleThrGlnAmL	-----	304
Qy	766	-----GGAACMAAGGST-----A	778	
Db	304	euserAspGlyLeuAsrAlaIalValSerGlyAlaIalagIyPheGlyLeuMetAlaTyv	-----	324
Qy	779	TCAGTTTGTCTCGTAGGCTTGGTATTTATCTCTCCATCAGTCATGTTGAAGATGC	838	
Db	339	IeGlyLeuGlyAlaPhePheTyrlArgAlaIleArgAsnLeuSerArgGlnIleGlnA	359	
Qy	839	TTGTGAGTGGTGTTTTGTATTAC-----AAATGCGGTGCCAGTTTGGTATGACCTTG	895	
Db	359	rgAsnLeuGlyThrLeuThr-----	365	
Qy	956	GAATGAATGAATGAACAACACATATATAGAGSTATGCTGTATAGCTACAGCTGTA	1015	
Db	366	-----LysIleAlaGlnGluArgLeuGlyA	374	
Qy	1016	ATTATGAACCTGAGATTTTATATATGAAGATGAAGATGAACAGACAGATATGATGAT	1075	
Db	374	snValLysThrSerGlnSerPheAlaGlyValLeuGluValArgArgyFrAmAsnG	394	
Qy	1076	TTTGTG---AAGACGTAGACACTGCTTCATTGAAAGTACCTCTACTGCTGATGCTGA	1132	
Db	394	InValArgLysIlePheGlu-----LeuGlyLysGlnSerLeuIleSerAlaT	411	
Qy	1133	ACTTGGTCMAAGGCTATTTTCAGTGTGGTTTACAGSTATAAGTGAGTGCGGACAGTC	1192	
Db	411	hrPhePheSerSerThrGlyPheAlaGlyAsnMetThrIleLeuAlaLeuLeuTyValG	431	
Qy	1193	AGGGA-----ATTGTGACAGTACCTCTTACTGTGGAGATCA-----	1230	
Db	431	lyGlyGlyMetValGlnSerIlyAlaIleThrIleGlyGluLeuThrSerPheLeuMet	451	
Qy	1231	-----GTAATGTGTAATGAGACTGCTTTTTCAGCTTTCATTATCCCTGAACTTTGGGA	1285	
Db	451	yrThrAlaTyAlaGlySerSerMetPheGlyLeuSer-----S	464	
Qy	1286	CTGTATATATGAGAGCTATGACAACTCATATAGATATGAACACTGTTTACTATATCA	1345	
Db	464	ePheTySerIleuLeuMetLysGlyAlaGlyAlaAserAsgLeuPheGluLeuGlnA	484	
Qy	1346	AGGTAGACACCCAAATTAAAGACAAAGTATGGCATCTCCCTT-----CAGATCA	1396	
Db	484	spArgGlnProThrIle-----SerProHrlysgGlyGluTyValA	498	
Qy	1397	CACCAACAGACACTACCGTGCCTTGATATATGATGCTATTGGAATC-----ATTGAGG	1450	
Db	498	IaSerAlaArgGlyProIleArgPheGluAsnValThrPheSerTyProThrArgProA	518	
Qy	1451	GCCAGAAAGTCTTATAGTGAATATCTTTGAAGTCCCTGAGAGAAAGAGGGCATTG	1510	
Db	518	IaValProIlePheArgAspLeuAsnPheGlnIleProGlnIlyThrAsnValAlaIleV	538	
Qy	1511	TAGAGATATAGTGGGACAGGAAAGACAAATATGTAAGAGCATATATTCGCTTCTATAGAC	1570	
Db	538	AlGlyProSerArgIlyGlyIlySerSerIleAlaSerIleLeuLeuArgPheTySerP	558	
Qy	1571	CTCAAAAGGGTATGACTTTATCTTGTGTCGTCATAATATCAAGATGTAGCTGGAAGCC	1630	
Db	558	roThrGlnGlyArgValLeuIleGlyLysAspIleThrIleMetAsnAlaLysSerL	578	
Qy	1631	TTTCGAGGGCAGTGGAGATGTTACTCAGAGATGCTGTCTTCCTTCATAATACTATTAT	1690	

[illegible]

ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5080
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1275 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-120-513-2

Alignment Scores:

Pred. No.:	5,61e-51	Length:	1275
Score:	583.50	Matches:	182
Percent Similarity:	46.65%	Conservative:	124
Best Local Similarity:	27.74%	Mismatches:	257
Query Match:	13.72%	Indels:	93
DB:	3	Gaps:	19

AF133659 (1-2345) x US-09-120-513-2 (1-1275)

```
QY 415 AGAGTTGCGATTGCGTGGGATTTTGGGTGCGCAAGGCCATGAATATTGTGTTCCC 474
D 46 LysLeuCyMetAlaLeuGlyThrLeu---AlaAlaIleIleHisGlyThrLeuLeuPro 64
QY 475 TTCATG-----TTTAATATGCTGTAGACAGCTTCAACGATGTCGGGAAACATG 525
D 65 LeuLeuMetLeuValPheGlyTyrMetThrAspSerPheThrGlnAlaGluThrArgIle 84
QY 526 CTG-----AACCTGAGTATGACCAAAATACAGTTGCAACCATGCGCAACAGCA 573
D 85 LeuProSerValThrAsnGlnSerGluIle---AsnSerThrGlnThrValSerAspSer 103
QY 574 GTTCTG-----ATTGCTATMGGTGTATCAAGA 600
D 104 SerLeuGlnGluAspMetAlaMetTyrAlaTyrTyrTyrThrGlyIleGly----- 120
QY 601 GCTGAGACTGCTTTTAAACGAAGTTGCAAGATTCAGATTGGCAAGTAGCCACAGAT 660
D 121 AlaeGlyValLeuIleValAlaTyrIleGlnValSerLeuTyrPyrLeuAlaIleGlyArg 140
QY 661 TCAATCCGAAGATAGCCAAAATGCTTCTCCATCTTCAACCTGATGCTGGGTTT 720
D 141 GlnIleHisTyrIleArgGlnLysPhePheHisAlaIleMetAsnGlnGluIleGlyTyr 160
QY 721 CACCTGAGCAGACAGACGGGAGCTTTATCTTAAGCTATGTACACAGAAACAAGGGTATC 780
D 161 PheAspValAsnAspAlaGlyGlnLeuAsn-----ThrArgLeuThrAspAspVal 177
QY 781 AGTTTCTCCTGAGTCTTGGATTATTAATCTTCCCATCATGTTGAAGTATGCTT 840
D 178 SerTyrIleAsnAspGlyIleGlyAspLysLeuGlyMetPhePheGlnSerIleThrThr 197
QY 841 GTCAGTGGT-----GTTTGTATTACAATGCGGTCGCCAGTTGCTTGGTAAACCTT 894
D 198 PheSerTyrIleGlyPheIleIleGlyPheIleSerGlyTyrPyrLysLeuThrLeuValIleLeu 217
QY 895 GGAACA-----CTTGATACATACACAGCATTCACAGTGCAGTCAGCACACGGTGAAGA 945
D 218 AlaValSerProLeuIleGlyLeuSerSerAlaMetTyrAlaLysValLeu----- 234
QY 946 ACTAGATTGATAGAAATGAACAAGACAGATATATGATGACAGTATGCTGCTATAGAC 1005
D 235 ThrSerPheThrAsnLysGlnLeuGlnAlaTyrTrpAlaLysAlaGlyAlaValAlaGluGln 254
QY 1006 TCACTGCTGAATTATTAAGAACTGTGAAGTATTTTATATATGAAGATATGAACAGAGA 1065
```

```
D 255 ValLeuAlaIleArgThrValIleAlaPheGlyGlyGlnLysGlnLeuGluArg 274
QY 1066 TATGATGATTTTGAAGACGTATGAGACTGCTTCATGAAAGTACCTACTGCT 1125
D 275 TyrAsnLysAsnLeuGlnGluAlaLysArgValGlyIleLysLysAlaIleThr----- 292
QY 1126 ATGCTGAACCTTTGGTCAAAAGTCTATTTTCAAGTGTGCGTTTAAACAGCTATTAAGTGTCTC 1185
D 293 -----AlaAsnIleSerIleGlyIleAlaTyrLeuLeuValTyr 305
QY 1186 GCCAGTCAG----- 1194
D 306 AlaSerTyrAlaLeuAlaPheTyrGlyThrSerLeuValLeuSerAsnGluTyrSer 325
QY 1195 -----GCAATTGTGGCAGTACCTCTTACTGTTGAGAT 1227
D 326 IleGlyGlnValLeuThrValPhePheSerIleLeuLeuGlyThrPheSerIleGlyHis 345
QY 1228 CTAGTAATGTTGAATGAGACTGCTTTTCAAGCTTTCATTAACCTGAACTTTCTGGGAAGT 1287
D 346 LeuAlaPro-----AsnIleGlnAlaPheAlaAsnAlaArgGlyAla 359
QY 1288 GTATATPAGAGACTAGACACAGCACTCATGATATGAACAACCTTGTTACTTACTGACAG 1347
D 360 AlaTyrGlnIlePheLys-----IleIleAspAsnGluProSerIleAspSerPheSer 377
QY 1348 GTAGACACCCAAATTTAAAGACAAAGATGATGCACTCCCTTCAGATCAACACACAGACA 1407
D 378 ThrLysGlyHisLysProAspSerIleMetGlyAsn----- 389
QY 1408 GCTACCGTGACCTTGATTAATGTGCATTTTGAATCATTTGAGGGCCAG-----AAAGTC 1461
D 390 -----LeuGlnPheLysAsnValTyrPheAsnTyrProSerArgSerGluValLysIle 407
QY 1462 CTTAGTGAATATACCTTTGAAGTCCCTGACGAAAGAAAGTGGCCATTTGAGAGATAGT 1521
D 408 LeuLysGlyLeuAsnLeuLysValLysSerGlyGlnThrValAlaLeuValGlyAsnSer 427
QY 1522 GGGTCAGGGAAGAAACCAATAGAGAGCTTATTTGCTTCTATGAGCTCAAAAGGCT 1581
D 428 GlyCyaeGlyLysSerThrThrValGlnLeuLeuGlnArgLeuTyrAspProIleGluGly 447
QY 1582 AGCATTTATCTGCTGCTCAAAATATACAGATGAGAGCTGGAAGACCTTGGAGGGCA 1641
D 448 GluValSerIleAspGlyGlnAspIleArgThrIleAsnValArgTyrLeuAspGluIle 467
QY 1642 GTGGAGTGTACTCAGATGCTGCTCTTCCATTAATCTTATTTTCAACCTCTTA 1701
D 468 IleGlyValValSerGlnGlnProValLeuPheAlaThrThrIleAlaGluAsnIleArg 487
QY 1702 TATGAAACATCAGTGCCTTCACTGAGGAAGTATGACAGTGCACAAATTAAGCTGACTT 1761
D 488 TyrGlyArgGlnAsnValThrMetAspGluIleGlnLysAlaValLysGluAlaAsnAla 507
QY 1762 CATGATGCAATTTCTTGAATGCCATGTATGACACCCAGTAGGGGAGAGAGATC 1821
D 508 TyrAspPheIleMetLysLeuProHisLysPheAsnThrLeuValGlyGluArgGlyAla 527
QY 1822 AACCTTTCAGGAGGAGAAAGCAAGATGACATTTGCAAGACCTTTGAAAGACCCC 1881
D 528 GlnLeuSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 547
QY 1882 CCACTCATCTATGATGAAGTACTTCACTGTTGATTCGATTCTGAAGAGACTATT 1941
D 548 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspThrGluSerGluAlaValVal 567
QY 1942 CTTGATGCCATGAAGATGTGTCAACACAGAACTTATTTTCTATTTGCACACAGATTG 2001
D 568 GlnAlaIleLeuAspLysAlaArgGlnGluArgThrThrIleValIleAlaHisLeuGlu 587
QY 2002 TCAACAGTGTTCATGACATGAATCATTTGCTTGATCAGGGTGAAGTGAAGCCGAACGT 2061
D 588 SerThrValArgAsnAlaAspValIleAlaGlyPheAspGlyGlyValIleValIleGluGln 607
```


Db 488 TyrGlyArgGluAsnValThrMetAspGluIleGlyValAlaValLeuGluAlaAsnAla 507
 Qy 1762 CATGATGCATTTCTTCGATGCGACATGATGACACCCAGTACGAGGAAAGGAGATCTC 1821
 Db 508 TyrAspPheIleMetLeuLeuProHisIlePheAsnThrLeuValGlyGluArgGlyAla 527
 Qy 1822 AAGCTTTAGAGAGAGAAAGAAAGAAAGTACGATTTGCAAGAGCCATTTTGAAGACCCC 1881
 Db 528 GluIleuSerGlyGlyGluIleGluIleArgIleAlaIleAlaArgAlaLeuValArgAsnPro 547
 Qy 1882 CCAGCATCTACTCTATGATGAGAGCTACTCTACCTGATGATGATGATGATGATGATGAT 1941
 Db 548 LysIleLeuLeuLeuLeuLeuGluAlaThrSerAlaLeuAspThrGluSerGlyAlaVal 567
 Qy 1942 CTGGTGCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2001
 Db 568 GluAlaIleuAspIlySalArgGluIleArgThrIleValIleAlaHisArgIleu 587
 Qy 2002 TCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2061
 Db 588 SerThrValArgAsnAlaAspValIleAlaGlyPheAspGlyGlyValIleValGluGln 607
 Qy 2062 GGATCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
 Db 608 GlyAsnHisGluGluIleuMetLysGluIleGlyIleThrLeuValIleuThr 626
 Qy 2116 CATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2175
 Db 627 GluThrArgGlyAsnGluIleGluProGlyAsnHisAlaIleArgIleSerAsp 645
 Qy 2176 AATATATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2223
 Db 646 ThrGlyAlaSerGluLeuThrSerGluIleuSerLysSerProLeuIle 661

RESULT 15
 US-08-784-649A-2
 ; Sequence 2, Application US/08784649A
 ; Patent No. 5830697
 ; GENERAL INFORMATION:
 ; APPLICANT: Sikic, Branimir I
 ; APPLICANT: Chen, Gang
 ; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
 ; TITLE OF INVENTION: CYCLOSPORIN MODULATION
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,649A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J
 ; REGISTRATION NUMBER: Reg. No. 5830697 36,677
 ; REFERENCE/DOCKET NUMBER: 06037/007001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-322-5070
 ; TELEFAX: 415-854-0875
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1279 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

; MOLECULE TYPE: peptide
 ; US-08-784-649A-2
 Alignment Scores:
 Pred. No.: 6,32e-51 Length: 1279
 Score: 583.00 Matches: 176
 Percent Similarity: 46.55% Conservative: 114
 Best Local Similarity: 28.25% Mismatches: 257
 Query Match: 13.70% Indels: 76
 DB: 2 Gaps: 16

AF133659 (1-2245) x US-08-784-649A-2 (1-1279)
 Qy 499 AGCCTCAACGATGTCGGGAAACATGCTGATGATGATGATGATGATGATGATGATGAT 558
 Db 84 AsnLeuIleuAspLeuMetSerAsnIleThrAsnArgSerAspIleAsnAspThrGlyPhe 103
 Qy 559 ACCATG-----GCAACAGCATTTCTGATGATGATGATGATGATGATGATGATGAT 603
 Db 104 PheMetAsnLeuGluGluIleuAspMetThrArgTyrAlaTyrTyrSerGlyIleGlyAla 123
 Qy 604 GGAAGTCCTTTTAAAGAGATTCGAATGCAAGTATTTGGCAAGTACGCCAGATTC 663
 Db 124 GlyValLeuValAlaAlaTyrIleGlnValSerPheTrpCysLeuAlaIleGlyArgGln 143
 Qy 664 ATCCGAAGATAGCCCAAAATGCTTCTTCATCTTCACAACTGATCGGTTTTCAC 723
 Db 144 IleHisIleArgIleGlyGlnPhePheHisAlaIleMetArgGlnGluIleGlyTrpPhe 163
 Qy 724 CTGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
 Db 164 AspValHisAspValGlyGluLeuAsnThrArgLeuThrAspAspValSerLysIleAsn 183
 Qy 784 TTTGTCCTGAGT-----GCTTGGTATTAATCTTCTCCATCATGTTTGAAGTG 834
 Db 184 GluValIleGlyAspIleGlyMetPhePheGlnSerMetAlaThrPhePheThrGly 203
 Qy 835 ATGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
 Db 204 PheIleValIle-----PheThrArgGlyTrpIleuThrLeuValIleLeu 219
 Qy 895 GGA-----ACACTGGTATACATACAGCATTCACAGTTCAGTTCAGTTCAGTTCAGT 945
 Db 220 AlaIleSerProValLeuGlyLeuSerAlaIleAlaIleAlaIleValIleLeu 236
 Qy 946 ACTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
 Db 237 SerSerPheThrAspIlyGluLeuLeuAlaTyrAlaLysAlaGlyAlaValAlaGluGlu 256
 Qy 1006 TCACTGCTGAATATGAAACTGGAAGTATTTAATATGAAAGATATGAAAGCAGAGAGA 1065
 Db 257 ValLeuAlaIleAlaIleArgThrValIleAlaPheGlyGlyIleuLysGluLeuGlnArg 276
 Qy 1066 TATGATGATTTTGAAGAGATGAGACTGCTTCAATGAAAGTACCTTACTGCTGCT 1125
 Db 277 TyrAsnLysAsnLeuGluIleAlaLysArgIleGlyIleLysLysAlaIleThr 294
 Qy 1126 ATGCTGAATCTTGTCAAAGTCTATTTTCAGTGGGTTTAAACAGTATTAAGTACGCTC 1185
 Db 295 -----AlaAsnIleSerIleGlyAlaAlaPheLeuLeuIleTyr 307
 Qy 1186 GCCAGTCAGGGA-----ATTGGCAGGTCACCTTACT 1218
 Db 308 AlaSerTyrAlaLeuAlaPheTrpTyrGlyThrThrLeuValIleuSerGlyIleTyrSer 327
 Qy 1219 GTTGAGATCTAGTAATGTGAATGAGCTGCTTTT----- 1254
 Db 328 IleGlyGlnValLeuThrValPheSerValIleuIleGlyAlaPheSerValGlyGlnAla 347
 Qy 1255 -----CAGTTTCATTAACCTGGAATCTTCTGGGAACGTATATAGAGAGACTAGACAA 1308
 Db 348 SerProSerIleGluAlaPheAlaAsnAlaArgGlyAlaAlaTyrGlu----- 363

```
OY 1309 GCACATCATAGATATGAAACCTTGTCTACTCTACTC-----AAGGTAGACACC 1356
Db 364 -----ilhehlysllelleasphnlyseProserillehspSer 376
OY 1357 CAAATTAAAGACAAAGTATGGCATCTCCCTTCAGATCACACACAGACACTACCGTG 1416
Db 377 TyserlysserGlyHsllyseProasphsnlleys-----GlyAsnLeu 391
OY 1417 GCCTTGAATATGTCATTTTAAATACATTTAGAGGCAG-----AAGTCTTAGTGGA 1470
Db 392 GlupheargasnvalahispheserTyProserArglyseGluvalylsilleuylsGly 411
OY 1471 ATATCCTTTGAAGTCCCTGCAGAAAGAAAGTGCATTTGTAGAGAGTAGTGGGTCAAGG 1530
Db 412 LeuAsnLeuylsvalGlnserGlyGlnthrValahaleuvalGlyAsnserGlyCysGly 431
OY 1531 AAAAGCACATATGTAGAGCTATTATTGCTTCTATGAGCTCTCAAAAAGGTAGCATTTAT 1590
Db 432 lysSerThrThrValGlnleuMetGlnArgleuTyAspProThrGlnGlyMetValSer 451
OY 1591 CTTCCTGGTCAAAATATATCAAGATGAGAGCTGAAAGCCTTCGAGGGGCAGTGGAGTG 1650
Db 452 ValasphGlyGlnasphllearghThrleasvalargpheleuargGluileilelyVal 471
OY 1651 GTACCTCAGAGATGCTGTCTCTTCATAATATCTATTATTACACCTCTTATATGGAAC 1710
Db 472 ValserGlnGluProValleuPhealThrThrillealaleuAsnilleargtyrGlyArg 491
OY 1711 ATCAGTCTTCACCTGAGGAAGTATGAGTGCGAAATTAGCTGAGCTTCATGATGCA 1770
Db 492 GluAsnValThrMetAspGlnleuGlnlysalavalylsGluAlaAsnAlaTyAspPhe 511
OY 1771 ATCTTCGAATGCACATGATATGACACCAAGTAGGGGAGACGAGACTCAAGCTTTCA 1830
Db 512 lleuMetlyseuProhlslysePheasphThrleuvalGlyGluargGlyAlaGlnleuSer 531
OY 1831 GGAGAGAAAGAAAGCAAGATAGCAATTGCAAGAGCCATTGAGAGACCCCGCAGTCATA 1890
Db 532 GlyGlyGlnlyslargllealalealalealaleuvalArgAsnProlysilleu 551
OY 1891 CTCTATGATGAAGCTACTTCATCGTTAGATTGAGATTCTGAAGAGACTATTCTTGTCGC 1950
Db 552 leuAspGlnAlaThrserAlaAsnThrGlnserGlnAlaValAlaGlnValAla 571
OY 1951 ATGAAGATGTGTGCAACACAGACTTCTATTTCATTCGACACACAGATTGCAACAGTG 2010
Db 572 leuAspIysAlaArglyseGlyArgThrThrillevalillealalehArgleuSerThrVal 591
OY 2011 GTTGATGCAGATGAATCATTTGTTGGATCAGGTAGTAGCCGAAACGTGTACCCAC 2070
Db 592 ArgAsnAlaAspValillealaglyPheAspAspGlyValillealGlnlyseGlyAsnHis 611
OY 2071 CATGTTTCCTTGCTTACCTCATAGTATCTATTCAGAAATGTGGCATACACAGACAGC 2130
Db 612 AspGlnleuMet---LysGlnlyseGlyIleTyPheylsleuvalThrMetGlnThrAla 630
OY 2131 -----CGTGTGCAGAAC---CATGATTAACCCCAATGGGAGCAAAAGAAAGAA 2175
Db 631 GlyAsnGluvalGlnleuGlnAsnAlaAlaAspGlnserLysSerGlnleAspAlaLeu 650
OY 2176 AATATATCCAAAGAGAGAAAGAAAGAACTA-----CAAGAGAAATTTGTCAATAGT 2229
Db 651 GluMetSerSerAsnAspSerArgserSerleuilearglyAsrSerThrArgArgSer 670
OY 2230 GTGAAAGGC 2238
Db 671 ValArgGly 673
```

Search completed: February 15, 2003, 06:10:04
Job time : 46.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 03:23:25 ; Search time 65 Seconds

(without alignments)
9614.545 Million cell updates/sec

Title: AF133659

Perfect score: 4254
Sequence: 1 ATGGCGCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DBV=xlh
-O=gsn2_1/USPFO.spool/AF133659/runar.10022003.155440.29503/app_query.fasta.1.2503
-DB=A.Geneseq.101002 -OPMT=factan -SIFPIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AF133659 @CGN 1.1 83 @runat 10022003.155440.29503 -NCPU=6 -ICPU=3
-NO XIPXY -NO MAP -TAG=QUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A. Geneseq.101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3809	89.5	747	20	Human ATP-binding
2	3809	89.5	747	21	Human ATP-binding
3	3736	87.8	752	20	Human ABC-Transpor
4	3656.5	86.0	722	22	Human protein sequ
5	1852	43.5	606	22	Drosophila melanog
6	1852	43.5	606	22	ABBS7811
7	1239.5	29.1	842	22	AAU09965
8	1239.5	29.1	842	23	AAE16764
9	1231	28.9	843	22	AAU00010
10	1218	28.6	843	22	AAU00011
11	1201.5	28.2	475	20	AAV08961
12	1183	27.8	574	21	AAV54454
13	1183	27.8	866	22	ABB71210
14	1174	27.6	571	22	AAV39313
15	1149.5	27.0	591	22	AAV41099
16	964.5	22.7	342	21	AAV17365
17	920.5	21.6	327	21	AAV17366
18	889.5	20.9	319	21	AAV17367
19	739	17.4	147	22	AAV75566
20	736.5	17.3	640	23	AAO22153
21	699.5	16.4	578	22	AAU36908
22	695.5	16.3	582	23	ABP39885
23	685	16.1	621	22	AAV92266
24	685	16.1	621	22	AAV76680
25	685	16.1	621	23	AAE22051
26	682	16.0	548	22	AAV81616
27	668	15.7	577	23	ABV47542
28	654	15.4	582	22	AAU38413
29	651	15.3	582	22	AAU34522
30	647.5	15.2	543	22	ABV47270
31	643	15.1	593	23	ABV49177
32	641.5	15.1	707	18	AAV22153
33	641.5	15.1	707	18	AAV22153
34	637.5	15.0	580	23	ABV40767
35	635	14.9	579	22	AAV85670
36	632	14.9	587	22	AAV35367
37	632	14.9	587	22	AAV35367
38	632	14.9	587	22	AAV35367
39	629.5	14.8	710	20	AAV27214
40	629.5	14.8	710	20	AAV27214
41	625	14.7	605	23	ABV47710
42	622	14.6	1279	22	ABV60234
43	621.5	14.6	583	22	AAV96741
44	621	14.6	1286	21	AAV39101
45	619	14.6	595	23	ABV25492

ALIGNMENTS

RESULT 1
ID AAV89585 standard; Protein; 747 AA.

AAV89585;
22-MAR-1999 (first entry)

Human ATP-binding cassette transport protein.

Human, ATP-binding cassette transport protein; ABCCxH; cancer;

neurological disorder; Alzheimer's disease; dementia; depression;

Down's syndrome; epilepsy.

Homo sapiens.

US5858719-A.
12-JAN-1999.

QY 1576 AAGGTAGCATTTATTTCTGCTGCTCAAAATATPACAAAGTGAAGCTGGAAAGCTTCGG 1635
DB 521 LysGlySerIleTyrIleuAlaGlyGlnAsnIleGlnAspValSerLeuGlnSerLeuArg 540
QY 1636 AGGGCAGTGGAGTGTACTGAGGATGCTGCTCTTCCATTAATCTTTATTTATTAAC 1695
DB 541 ArgAlaValGlyValValProGlnAspAlaValLeuPheHisAsnThrIleTyrIleAsn 560
QY 1696 CTCTTATATGAAACATCAGTCTTCACCTGAGAGAGTATGACAGTGGCAAAATAGCT 1755
DB 561 LeuLeuTyrGlyAsnIleSerIleSerProGlnGlnValTyrAlaValAlaLysLeuAla 580
QY 1756 GCACTTCATGATGCAATTTCTTGCAATGCCACATGATATGACACCCAGTAGGGAGCA 1815
DB 581 GlyLeuHisAspAlaIleLeuArgMetProHisGlyTyrAspThrGlnValGlyIleArg 600
QY 1816 GGACCTCAAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875
DB 601 GlyLeuLysLeuSerGlyGlyGlnLysGlnArgValAlaIleAlaArgAlaIleLeuLys 620
QY 1876 GACCCCCAGTCATCTCTATGATGAGAGTCTTCATCGTTAGATTGCAATTAAGAG 1935
DB 621 AspProProValIleLeuTyrIleAspGlnAlaThrSerSerLeuAspSerIleThrGln 640
QY 1936 ACTATTTCTGTCGCAATGAGAGATGTGTCAACAACAGAACTTCTATTTTCATTCGAC 1995
DB 641 ThrIleLeuGlyAlaMetLysAspValValLysHisArgThrSerIlePheIleAlaHis 660
QY 1996 AATTTGTCAACGTGTGTGATGATGCAATGAAATCATTTGTTTGGATCAGGGTAAAGTACC 2055
DB 661 ArgLeuSerThrValValAlaAspAlaAspGlnIleIleValLeuAspGlnIleLysValAla 680
QY 2056 GAACGTGTGACCCACATGCTGTTGCTGCTGTAACCTCATATGATATTCATCAAAATGTGG 2115
DB 681 GlnArgGlyThrHisHisGlyLeuLeuAlaAsnProHisSerIleTyrSerIleuMetTrp 700
QY 2116 CATACACAGAGAGCCGTGTGCAAGAACCATGATTAACCCCAATGGAGCAAAAGAGAA 2175
DB 701 HistHrGlnSerSerArgValGlnAsnHisAspAsnProLysTrpGlnAlaLysGln 720
QY 2176 AATATATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2235
DB 721 AsnIleSerLysGlnGlnGlnIleArgLysLysLeuGlnGlnIleValAlaAsnSerValLys 740
QY 2236 GGCTGTGAAAAGCTGTTGCTGC 2256
DB 741 GlyCysGlyAsnCysSerCys 747
RESULT 2
AAB13355
ID AAB13355 standard; Protein; 747 AA.
XX AAB13355;
AC
XX
DT 13-NOV-2000 (first entry)
XX
DE Human ATP-binding cassette transport protein.
XX
KW Human; ATP-binding cassette transport protein; ABCtxH; cytosolic;
KW neurotrophic; neuroprotective; cerebroprotective; antidepressant;
KW anticonvulsant; antiparkinsonian; neuroleptic; cancer; epilepsy;
KW neurodegenerative disorder; akathisia; amyotrophic lateral sclerosis;
KW bipolar disorder; cataplexia; dementia; depression; Down's syndrome;
KW tardive dyskinesia; dystonia; multiple sclerosis; neurofibromatosis;
KW schizophrenia; Tourette's disease.
XX
OS Homo sapiens.
XX
PN US608042-A.
XX
PD 27-JUN-2000.
XX
PF 18-NOV-1998; 98US-0195391.

XX 17-JUL-1997; 97US-0895522.
XX
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Shah P, Corley NC, Hillman JL;
XX WPI; 2000-451228/39.
DR N-PSDB; AAA62411.
XX
PT New human ATP-binding cassette transport protein useful for diagnosing,
PT preventing and treating cancers, e.g. leukemia, sarcoma, and neuronal
PT disorders, e.g. schizophrenia, dementia
XX
PS Claim 1; Fig 1; 31pp; English.
XX
CC The present sequence is a human ATP-binding cassette transport protein
CC (ABCTxH). The nucleotide sequence encoding ABCTxH was isolated from a
CC human ovarian tissue cDNA library. Clones from the library were sequenced
CC and used as query sequences against homology databases. ABCTxH encoding
CC polynucleotides were extended using PCR. ABCTxH polynucleotides and
CC polypeptides may be used for the diagnosis, prevention of ABCTxH, e.g.
CC conditions or disorders associated with the expression of ABCTxH, e.g.
CC adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma,
CC teratocarcinoma, akathisia, Alzheimer's disease, amnesia, amyotrophic
CC lateral sclerosis, bipolar disorder, cataplexia, cerebral neoplasms,
CC dementia, depression, Down's syndrome, tardive dyskinesia, dystonias,
CC epilepsy, Huntington's disease, multiple sclerosis, neurofibromatosis,
CC Parkinson's disease, paranoid psychoses, schizophrenia, and Tourette's
CC disease.
XX
SQ Sequence 747 AA;
Alignment Scores:
Pred. No.: 0 Length: 747
Score: 3809.00 Matches: 746
Percent Similarity: 99.87% Conservative: 0
Best local Similarity: 99.87% Mismatches: 1
Query Match: 89.54% Indels: 0
DB: 21 Gaps: 0
AF133659 (1-2345) x AAB13355 (1-747)
QY 16 ATGCATCTTGCGCGTGGCGCGCGCGCGCTGCTTCAAAAAGCGCGGACATCCGCG 75
DB 1 MethSerTrpArgTrpAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 20
QY 76 ATTCGATCCGCGCTTATAGTCTCTGTTAGCGGCTCAGGTCGCGAGTGGAGGCCATCA 135
DB 21 IleuIleArgProLeuValSerValSerGlySerGlyProGlnTrpArgProHisGln 40
QY 136 CTGGCGCGCTTGGAGACCGCTCGAGCTTACCAATTCAGAGCTTAAAGATTCACA 195
DB 41 LeuGlyAlaLeuGlyThrAlaArgAlaTyrGlnIleProGlnSerLeuLysSerIleThr 60
QY 196 TGGCAGAGATTGGGAAAAGCAATTCAGAGAGTCTTATAGTGGCAAAAGCTTCAG 255
DB 61 TrpGlnArgLeuGlyLysGlyAsnSerGlyGlnPheLeuAspAlaAlaLysAlaLeuGln 80
QY 256 GTATGGCCATGATAGAAAAGAGACATGTTGGCATGGTCAAGAGAGAGACTCCAC 315
DB 81 ValTrpProLeuIleGlyLysArgThrCysTrpHisGlyHisAlaGlyGlyGlyLeuHis 100
QY 316 ACAAGCCAAAAGAGGTTAAAGATGTGATCTCGGAAAATCATTAAGCAATGCTT 375
DB 101 ThrAspProLysGlnGlyLeuLysAspValAspThrArgLysIleIleLysAlaLysLeu 120
QY 376 TCTATGATGAGCCCAAGACAGCCAGATTCAGAGTAGAGTTGCCATTTGCTGGGA 435
DB 121 SerTyrValTrpProLysAspArgProAspLeuArgAlaArgValAlaIleSerLeuGly 140
QY 436 TTTTGGGTGTGCAAGGCCATGAATATGTGCTCCCTTCATGTTAAATATGCTGTA 495

Db 141 PheLeuGIyGIyAlaLysAlaMetAenIleValIleProPheMetPheUySryAlaVal 160
 QY 496 GACAGCTCAACCAAGATGTCGGGAAACATGCTGACCTGATGCACCAATTAACGTT 555
 Db 161 AspSerLeuasnGlnMetSerGIyAsnMetLeuasnLeuSerAspAlaProAsnThrVal 180
 QY 556 GCACCAATGGCAACACAGCTTCTGATTTGGCTATGGTATCAAGACCTGGAGCTGCTTTT 615
 Db 181 AlaThrMetAlaThrAlaValLeuIleGIyTyrGIyValSerArgIleGIyAlaAlaPhe 200
 QY 616 TTTTAACGAAGTTGAAATGACGATTTGGCAAGCTAGCCCAATTCATCCGAAGAAT 675
 Db 201 PheAsnGIyValArgAsnAlaValPheGIyLysValAlaGlnAsnSerIleAspArgIle 220
 QY 676 GCCAAATAATGCTTCTCCATCTTCAACCTGATCTGGGCTTTTCACTGACGACAG 735
 Db 221 AlaLysAsnValPheLeuHISLeuHISAsnLeuAspLeuGIyPheHISLeuSerArgGln 240
 QY 736 ACCGAGCTTTATCTAAGCTATTGACAGAGAAAGGGATATCAGTTTGTCTGAGT 795
 Db 241 ThrGIyAlaLeuSerIyAsnIleAspArgGIyThrArgGIyIleSerPheValLeuSer 260
 QY 796 GCTTTGGATTTAATCTTCTCCATCATGTTGAAAGTATGCTTGTCACTGGTGTGTTTG 855
 Db 261 AlaLeuValPheAsnLeuLeuProIleMetPheGIyValMetLeuValSerGIyValLeu 280
 QY 856 TTTTAATAATGGGTCGCCAGTTTGGTTGGTAAACCTTGGACACTGGTATACATCA 915
 Db 281 TyrTyrLysSerGIyAlaGlnPheAlaLeuValThrLeuGIyThrLeuGIyThrTyr 300
 QY 916 GCATTCAAGTTGACAGTCAACCGGTGAGAGAACTGATTTGAATAGAAATGAACAAAGCA 975
 Db 301 AlaPheThrValAlaValAlaThrArgTyrArgThrArgPheArgIleGIyMetAsnLysAla 320
 QY 976 GATATGATGACAGTAAATGCTGCTATAGACTCACTGCTGATTAATGAACCTGGAAGTAT 1035
 Db 321 AspAsnAspAlaGIyAsnAlaAlaIleAspSerLeuAsnTyrGIyThrValIySry 340
 QY 1036 TTTTAATAAGAAATGATAGAGACAGAGATATGATGATTTTGAAGCTGATGAGCT 1095
 Db 341 PheAsnAsnGIyLysArgIyLysAlaGlnArgTyrAspGIyPheLeuIySrnTyrGIyLys 360
 QY 1096 GCTTCATGAAAAGTACCTTACTGCTGCTATGCTGATGCTGAACTTGGTCAAAAGTCTATTTC 1155
 Db 361 AlaSerLeuIySerThrSerThrLeuAlaMetLeuAsnPheGIyGlnSerAlaIlePhe 380
 QY 1156 AGTGTGGTTTAAACAGCTATATGCTGCTGCCAGTCAAGGAATGTGGCAGGTACCTT 1215
 Db 381 SerValGIyLeuThrAlaIleMetValLeuAlaSerGlnGIyIleValAlaGIyThrLeu 400
 QY 1216 ACTGTTGGAGATCTAGTAATGGAATGACCTGTTTTCAGTTTCATTACCCCTGAAC 1275
 Db 401 ThrValGIyAspLeuValMetValAsnGIyLeuLeuPheGlnLeuSerLeuProLeuAsn 420
 QY 1276 TTTCTGGAACTGTATATAGAGACTAGACAAGCACTCATAGATATGAACCTTGT 1335
 Db 421 PheLeuGIyThrValIyTyrArgGIyThrArgGlnAlaLeuIleAspPheAsnThrLeuPhe 440
 QY 1336 ACTCTACTCAAGGTAGACACCCAAATTAAGACAAAGTGGCARTCTCCCTTCAGATC 1395
 Db 441 ThrLeuLeuIyValAspThrGlnIleLysAspLysValMetAlaSerProLeuGlnIle 460
 QY 1396 AACCAACAGACACTACCGTGGCTTGTGAATGTGATTTTGAATATCATATGAGGGCCAG 1455
 Db 461 ThrProGlnThrAlaThrValAlaPheAspAsnValHISpHeGIyTyrIleGIyGlnGIy 480
 QY 1456 AAAGTCTTAGTGAATATCTTTGAAATCCCTGCAGAGAAAGAAAGGCGCATTTGAGA 1515
 Db 481 LysValLeuSerGIyIleSerPheGIyValProIleGIyLysValAlaAlaIleValGIy 500
 QY 1516 GGTAGTGGGTCAAGGAAAGCAACATATGAGGCTATATTTGCTTCAAGACCTCA 1575
 Db 501 GlySerGIySerGIyLysSerThrIleValArgLeuLeuPheArgPheTyrGIyLysProGln 520

QY 1576 AAGGTAGCATTTATCTTGCTGTGCAAAATATACAGAATGTGAGCCTGGAAACCTTCGG 1635
 Db 521 LysGIySerIleTyrLeuAlaGIyGlnAsnIleGlnAspValSerLeuGlnSerLeuArg 540
 QY 1636 AGGGCAGTGGAGGTGTATCTCAGATGTGTCTCTTCATATATCTATTTATTATCAAC 1695
 Db 541 ArgAlaValAlaGIyValIleProGlnAspAlaValLeuPheHISAsnThrIleTyrTyrAsn 560
 QY 1696 CTCTATATNGAACAATCAAGTCTTCACTGAGGAAGGTATGACACCCCAATGAGTAACTAGCT 1755
 Db 561 LeuLeuTyrGIyAsnIleSerAlaSerProGlnGIyValTyrAlaValAlaLysLeuAla 580
 QY 1756 GGACTTCATGATGACAAATCTTCTGAAATGCCACATGATATGACACCCCAATGAGGAACGA 1815
 Db 581 GlyLeuHISAspAlaIleLeuArgMetProHISGIyTyrAspThrGlnValGIyLysArg 600
 QY 1816 GGACTCAAGCTTTCAAGAGAGAGAAAGCAAGAGTATGCAATTTGCAAGCCATTTTGAAG 1875
 Db 601 GlyLeuLysLeuSerGIyGIyLysGlnArgValAlaIleAlaArgAlaIleLeuLys 620
 QY 1876 GACCCCGCATCTCTATGATGAGCTACTGATGATGATGATGATGATGATGATGATGATGAT 1935
 Db 621 AspProIleValIleLeuTyrAspGIyLysAlaThrSerLeuAspSerIleThrGIyGln 640
 QY 1936 ACTATTTCTGTGTCATGAGAGATGTGTCAACAACAACTTCTAATTTTCATTGCAAC 1995
 Db 641 ThrIleLeuGIyAlaMetLysAspValValLysHISArgThrSerIlePheIleAlaHIS 660
 QY 1996 AGATTGTCAACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
 Db 661 ArgLeuSerThrValValAlaAspAlaAspGlnIleIleValLeuAspGlnGIySerValAla 680
 QY 2056 GAACGTGTGACCAACAGTGTGTTGCTGTGTAACCTCATAGATGATGATGATGATGATGAT 2115
 Db 681 GluArgGIyThrHISIleGIyLeuLeuAlaAsnProHISerIleTyrSerGlnMetCtp 700
 QY 2116 CATACACAGACAGCGGTGTGAGAACCATGATTAACCCCAATGGGAGCAAGAAAGAA 2175
 Db 701 HISThrGlnSerSerArgValGlnAsnHISAspAsnProLysTyrGIyLysGIy 720
 QY 2176 AATATATCCAAAGAGAGAGAAAGAAAGAACTTACAAAGAAATTTGCTAATGTGTGAAA 2235
 Db 721 AsnIleSerLysGIyGlnGIyLysArgLysLysLeuGlnGIyIleValAsnSerValLys 740
 QY 2236 GGCTGTGAAACTGTCTGTC 2256
 Db 741 GlyCysGIyAsnCysSerCys 747
 RESULT 3
 AA014068
 ID AA014068 standard; Protein; 752 AA.
 AC AA014068;
 AC AA014068;
 DT 16-JUL-1999 (first entry)
 DT 16-JUL-1999 (first entry)
 XX Human ABC-Transporter-7 (HABCT), protein sequence.
 DE Human ABC-Transporter-7 (HABCT), protein sequence.
 XX HABCT; human; ABC-transporter-7; diagnosis; cancer; autoimmune disease;
 KW Addison's disease; insulin-dependent diabetes mellitus; therapy;
 KW microsome disorder.
 OS Homo sapiens.
 OS Homo sapiens.
 PN WO9921885-A1.
 PN WO9921885-A1.
 PD 06-MAY-1999.
 PD 06-MAY-1999.
 XX 29-OCT-1997; 97WO-CN00120.
 XX 29-OCT-1997; 97WO-CN00120.
 PR 29-OCT-1997; 97WO-CN00120.
 XX

PA (UYSH-) UNIV SHANGHAI SECOND MEDICAL.

XX Zhang Q;

XX WPI: 1999-312945/26.

DR N-PSDB; AAX36895.

XX HABC7 a protein useful in the treatment of cancer, diabetes,
PT microsomal disorders and Addison's disease

PS Claim 11; Page 27-29; 31pp; English.

XX This sequence is the human ABC-transporter-7 (HABC7) protein of the
CC invention. HABC7 is an ABC transporter protein. The DNA, vectors
CC containing it and recombinant host cells are useful for recombinant
CC production of HABC7. The DNA, HABC7 and antibodies against HABC7 are
CC useful as research reagents, for screening assays and in diagnostic
CC assays. Antagonists and agonists of HABC7 can be used to inhibit or
CC enhance, respectively, the activity of HABC7 or expression of the HABC7
CC coding sequence. Anti-HABC7 antibodies and HABC7 or its fragments can be
CC used in vaccines. In particular, the proteins, antibodies, agonists and
CC antagonists can be used for treating, e.g. cancer, autoimmune disease,
CC Addison's disease, microsomal disorders and insulin-dependent diabetes.
CC mellitus, related to both an excess and insufficient amounts of HABC7.

XX Sequence 752 AA;

Alignment Scores:

Pred. No.:	0	Length:	752
Score:	3736.00	Matches:	738
Percent Similarity:	98.54%	Conservative:	4
Best Local Similarity:	98.01%	Mismatches:	9
Query Match:	87.82%	Indels:	2
DB:	20	Gaps:	2

AF133659 (1-2345) x AAY14068 (1-752)

```

QY 1 ATGCGCTGCTCGCGATGATCTTTGGCGTGGCGCGCGCGCTGCTTTGCAAAAG 60
   |||||
DB 1 MetLaleuLeuAlaMetHisSerTtrpArgTtrpAlaAlaAlaAlaPheGluLys 20
QY 61 CGCGGCACTCCGCGATTCGATCCGCGCTTCTGCTTGAAGCGCTCAAGTCCGAG 120
   |||||
DB 21 ArgTrpHisSerAlaIleLeuIleArgProLeuValSerValSerGlyProGln 40
QY 121 TGGAGGCCACATCACTCGGCGCTTGGGAACCGTCAGCGCTACAGATTCGAGTCA 180
   |||||
DB 41 TtrpArgProHisGlnLeuGlyAlaLeuGlyThrAlaArgAlaTyGlnIleProGlnSer 60
QY 181 TTAAGAAGTATCACATGCGAGATTGGGAAAGGCAATTCAGGACAGTCTTGATGCT 240
   |||||
DB 61 LeuLysSerIleThrTrpGlnArgLeuGlyLysGlnSerGlyGlnPheLeuAspAla 80
QY 241 GCAAGGCTCTCAAGTATGCGCACTGATAGAAAAGAGACATGTTGGCATGTCATGCA 300
   |||||
DB 81 AlaLysAlaLeuGlnValTrpProLeuIleGluLysArgThrIleGlyHisAla 100
QY 301 GGAGAGGACTCCACACAGACCCCAAGAGGCTTAAAGATGTTGATCTCGAAAAATC 360
   |||||
DB 101 GlyIleGlyLeuHisThrAspProLysGlnIleLeuLysAspValAspThrArgLysIle 120
QY 361 ATAAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGGCGCAGATCTTCAAGCTAGATT 420
   |||||
DB 121 IleLysAlaMetLeuSerTrpValTrpProLysAspArgProAspLeuArgAlaArgAla 140
QY 421 GCCATTTGCTGGGATTTTGGTGTGTGCAAAAGGCGCATGAATATTTGGTTCCTTCATG 480
   |||||
DB 141 ProIleSerLeuGlyPheLeuGlyGlyAlaLysAlaMetCnIleValValProPheMet 160
QY 481 TTTAAATATGCTATGACAGCTCAACAGATGTCGGAAACAATGCTGAACCTGAGTAT 540
   |||||
DB 161 PheLysTrpAlaValAspSerLeuAsnGlnMetSerGlyAsnMetLeuAsnLeuSerAsp 180

```

```

QY 541 GCACCAATATACATGTCACACCATGGCAACAGACCTGCTGATGGCTATGCTATCAGA 600
   |||||
DB 181 AlaProSerThrValAlaThrMetAlaThrAlaValLeuIleGlyTyGlyValSerArg 200
QY 601 GCTGGAGCTGCTTTTAAAGAGTTGCAATGCAAGTATTTGGCAGGTAGCCCAAGAT 660
   |||||
DB 201 AlaGlyAlaAlaPhePheAsnGlnValArgAsnAlaValPheGlyLysValAlaGlnAsn 220
QY 661 TCAATCCGAAGAATAGCCAAATAATGCTTCTTCATGCTTCAACATCGATCGATGGATT 720
   |||||
DB 221 SerIleArgArgIleAlaLysAsnValPheLeuHisLeuHisAsnLeuAspLeuGlyPhe 240
QY 721 CACCTGACAGACAGACAGCGGAGCTTATCTAAGGCTATTCAGACAGAAACAGGGTATC 780
   |||||
DB 241 HisLeuSerArgGlnThrGlyAlaLeuSerLysAlaIleAspArgGlyThrArgGlyIle 260
QY 781 AGTTTGTCCGAGTGGCTTTGGTATTTAATCTTTCCCATCATGTTGAAGTGATG--- 837
   |||||
DB 261 SerPheValLeuSerAlaLeuValPheAsnProLeuProAsnHisValGluValMetLeu 280
QY 838 CTGTGACGTGGTGTGTTGATTAACAAATGCGGCGCCAGTTGCTTGGTAAACCTTGGA 897
   |||||
DB 281 LeuValSerGlyValLeuTrpTrpLysCysAlaGlnLeu---LeuGlyAsnLeuGly 299
QY 898 ACACCTGATCATACACAGCATTCACAGTTGCAGCTCACCGGTGAGAACTAGATTAGA 957
   |||||
DB 300 ThrLeuGlyThrTrpThrAlaPheThrValAlaValThrArgTrpArgThrArgPheArg 319
QY 958 ATAGAAATGAACAAAGACATATGATGACAGTATGCTGCTATAGACTCACTGCTGAAT 1017
   |||||
DB 320 LeuGlnIleAspGlnAlaAspAsnAspAlaGlyAsnAlaAlaIleAspSerLeuLeuAsn 339
QY 1018 TATGAACCTGGAAGTATTTAATGAAGAAGATGATAGACAGACAGATATAGATTT 1077
   |||||
DB 340 TyrluThrValLysTrpPheAsnAsnGlnArgTrpGlnAlaGlnArgTrpAspGlyPhe 359
QY 1078 TTGAAGACGTATGAGACTGCTTCAATGAAGAAGTCTTACTGCTGCTATGCTCAACTTT 1137
   |||||
DB 360 LeuLysThrTrpGlnThrAlaSerLeuLysSerThrSerThrLeuAlaMetLeuAsnPhe 379
QY 1138 GGTCAAGTGTATTTTCAAGTGTGCTTTAAACAGCTATATATGCTGCGCAGTCAAGGA 1197
   |||||
DB 380 GlyLysSerAlaIlePheSerValGlyLeuThrAlaIleMetValLeuAlaSerGlnGly 399
QY 1198 ATTGGGACAGGTACCTTACTGTTGAGATCTGATATGAGATGAGTCACTTTTCAG 1257
   |||||
DB 400 IleValAlaGlyThrLeuThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGln 419
QY 1258 CTTTCATTTACCCCTGAACCTTCTGGGAACGTATATAGAGAGACTAGACAAGACTCATA 1317
   |||||
DB 420 LeuSerLeuProLeuAsnPheLeuGlyThrValTyrluGlnThrArgGlnAlaLeuIle 439
QY 1318 GATATGAACACCTTGTACTTATCTCAAGGTAGACACCCAAATTAAGCAAAAGTAGT 1377
   |||||
DB 440 AspMetAsnThrLeuPheThrLeuLeuLysValAspThrGlnIleLysAspLysValMet 459
QY 1378 GCATCTCCCTTCAGATCACACACACAGACTACCGTGGCTTGTATATGTCATTTT 1437
   |||||
DB 460 AlaSerProLeuGlnIleThrProGlnThrAlaThrValAlaPheAspAsnValHisPhe 479
QY 1438 GAATACATTAAGGGCCAGAAAGTCTTAGTGAATATCTTTGAAGTCCCTGACAGAAAG 1497
   |||||
DB 480 GluTrpIleGluGlnLysValLeuSerGlyLysSerPheGluValProAlaGlyLys 499
QY 1498 AAAGTGCATTTGTAGAGGTAGTGGTCAAGGAAAGACAAATAGTAGAGCTTATTT 1557
   |||||
DB 500 LysValAlaIleValGlyLysSerGlySerGlyLysSerThrIleValArgLeuLeuPhe 519
QY 1558 CGCTTCATAGAGCTCAAAAGGATGATTTATCTTCTGCTGCTCAAAATATACAGAGTGT 1617
   |||||
DB 520 ArgPheTrpGlnProGlnLysGlySerIleTyrluValGlnAlaIleGlnAspVal 539
QY 1618 AGCCTGGAAGACCTTCGAGAGGCGAGTGGAGTGATCCTCAGATGCTGTCCTTCAT 1677

```

```

Db      540 SerLeuGluSerLeuArgArgAlaValAlaValProGlnAspAlaValLeuPheHis 559
      |||
Qy      1678 AATACTATTATTATCAACCTCTTATATGGAACATCAGCTTCCCTGAGGAAGTAT 1737
      |||
Db      560 AsnThrIleTyrTrpAsnLeuLeuTyrGlyAsnIleSerAlaSerProGlnIleValTyr 579
      |||
Qy      1728 GCAGTGGCAAAATTTAGTGCATTCATGATGCAATTTCTTGAATGCCATGATGATGAC 1797
      |||
Db      560 AlaValAlaLysLeuAlaGlyLeuHisAspAlaIleLeuArgMetProHisGlyTyrAsp 599
      |||
Qy      1798 ACCCAAGTAGGGGAAAGAGAGACTCAAGCTTTGAGAGAGAGAAAGCAAGATGCAAT 1857
      |||
Db      600 ThrGlnValGlyGluArgGlyLeuLysLeuSerGlyGlyGlnArgValAlaIle 619
      |||
Qy      1858 GCAAGAGCATTTTGAAGAGACCCCGAGTCATCTATGATGAAGAGCTACTTCATCTGTA 1917
      |||
Db      620 AlaArgAlaIleLeuLysAspProProValIleLeuTyrAspGluAlaThrSerSerLeu 639
      |||
Qy      1918 GATTGATTACTGAAGAGACTATTCTGTGGCATGAAGAGATGGTCAAAACAGAACT 1977
      |||
Db      640 AspSerIleThrGlnGluThrIleLeuGlyAlaMetLysAspValValLysHisArgThr 659
      |||
Qy      1978 TCTATTTTCATTGCACACAGATTGTCAACAGTGTGATGACAGATGAATCATTTGTTG 2037
      |||
Db      660 SerIlePheIleAlaHisArgLeuSerThrValAlaAspGluIleIleValLeu 679
      |||
Qy      2038 GATCAGGGTAGAGTACCCAGACAGTGGTACCACCATGTTGCTTCTGCTACCTCATACT 2097
      |||
Db      680 ArgGlnGlyLysValAlaGluArgGlyThrIleHisGlyLeuLeuAlaAsnProHisSer 699
      |||
Qy      2098 ATCTATTGCAATGTGCGATACACAGACAGCCGTTGTCAGAACCATGATTAACCCCAA 2157
      |||
Db      700 IleTyrSerGluMetIleThrHisThrGlnSerSerArgValGlnAsnHisAspAsnProLys 719
      |||
Qy      2158 TGGGAAGCAAGAAAGAAATATATTCCAAGAGAGAGAGAAAGAAACTACAGAAAGA 2217
      |||
Db      720 TrpGluAlaLysLysGluAsnIleSerLysGlnGluArgLysLeuGlnGluGln 739
      |||
Qy      2218 ATTGTCAATATGTGTGAAGAGCTGTGTGAACCTGTCTGTC 2256
      |||
Db      740 IleValAsnSerValLysGlyCysGlyAsnCySerCys 752
      |||

```

RESULT 4

AAB92941
ID AAB92941 standard; Protein; 722 AA.

AC AAB92941;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11610.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

OS Homo sapiens.

PM EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000BP-0116126.

PR 29-JUL-1999; 99TP-0248036.

PR 27-AUG-1999; 990P-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

```

XX      WPI; 2001-318749/34.
DR      Primer sets for synthesizing polynucleotides, particularly the 5602
XX      full-length cDNAs defined in the specification, and for the detection
PT      and/or diagnosis of the abnormality of the proteins encoded by the
PT      full-length cDNAs -
XX      Claim 8; SEQ ID 11610; 2537bp + CD ROM; English.
PS

```

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 722 AA;

Alignment Scores:

Pred. No.: 0 Length: 722
Score: 3656.50 Matches: 720
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 85.95% Indels: 1
DB: 22 Gaps: 1

AF133659 (1-2345) x AAB92941 (1-722)

```

Qy      1 ATGGGGCTGCTCGCATGATCTTGGCGGCGCGCGGCGGCTTGGAAAG 60
      |||
Db      1 MetAlaLeuLeuAlaMetHisSerTyrArgTrpAlaAlaAlaAlaPheGluLys 20
      |||
Qy      61 GCGCGCACTCCGCGATTCTGATCCGCGCTTACTCTCTGTTAGCGCGCTCAGTCCGAG 120
      |||
Db      21 ArgArgHisSerAlaIleLeuIleArgProLeuValSerValSerGlySerGlyProGln 40
      |||
Qy      121 TGGAGGCCATCAACTGGCGGCTTGGGAACCGCTCAGACCTTAC--CAGATTCAGAG 177
      |||
Db      41 TrpArgProHisGlnLeuGlyAlaLeuGlyThrAlaArgAlaTyrGlnGlnIleProGln 60
      |||
Qy      178 TCATTAAAGATATACATGCGAGATTTGGGAAGCAATTCAGACAGTCTTAGAT 237
      |||
Db      61 SerLeuLysSerIleThrTrpGlnArgLeuGlyLysGlyAsnSerGlyGlnPheLeuAsp 80
      |||
Qy      238 GCTGCAAGGCTCTCCAGGTATGGCCACTGATGAGAAAGAGAGACATGTTGGCATGTCAT 297
      |||
Db      81 AlaAlaLysAlaLeuGlnValTrpProLeuIleGluLysArgThrCysThrPheGlyHis 100
      |||
Qy      298 GCAGAGAGAGAGCTCCACACAGACCCCAAGAGGCTTAAAGATGTTGATCTCGGAAA 357
      |||
Db      101 AlaGlyGlyGlyLeuHisThrAspProLysGlyGlyLeuLysAspValAlaAspThrArgLys 120
      |||
Qy      358 ATCATTAAGCAATGCTTTCTTATATGTGTGGCCCAAGACAGGCGCATCTAGAGACTGA 417
      |||
Db      121 IleIleLysAlaMetLeuSerTyrValTrpProLysAspArgProAspLeuArgAlaArg 140
      |||

```


Qy 418 GTTGCCATTGCTGGAGATTTTGGGTGGTCAAAAGCCATGAATATGTGGTCCCTTC 477
Db 141 ValAlaIleSerLeuGlyPheLeuGlyGlyAlaIysSalIaMetAsnIleValValProPhe 160
Qy 478 ATGTTAAATATGCTGTAGACAGCCTCAACAGATGTGGGAAACATGCTGAACCTGAGT 537
Db 161 MetPheIyTrIyAlaValAspSerIeuAsnGlnMetSerGlyAsnMetIeuAsnIeuSer 180
Qy 538 GATGCACCAAAATAGACGTTGCAACCAATGGCAACAGCAGTTCGTATGGCTATGGTATCA 597
Db 181 AspIaIaProAsnThrValAlaThrMetAlaThrAlaValIeuIleGlyTyrGlyValSer 200
Qy 598 AGAGCTGGAGCTGCTTTTATTAAGCAAGTTGAAATGCAGATTTGGCAAGTACGCCAG 657
Db 201 ArgAlaGlyAlaIaIaPhePheAsnGlyValAlaArgAsnAlaValIlePheGlyLysValAlaGln 220
Qy 658 AATTCAATCCGAGAGATAGCCAAAAATGCTCTTCCATCTTCCCAACCTGGATCTGGCT 717
Db 221 AsnSerIleArgAlaGlyIleAlaIysAsnValIlePheIeuHISleuHISAsnIleuAspLeuGly 240
Qy 718 TTTGACCTGACAGACAGACCGGAGCTTTATCTAAGCTATGACAGAGAAAGGGGT 777
Db 241 PheIleIeuSerArgGlnThrGlyAlaIeuSerIySalIleAspArgGlyThrArgGly 260
Qy 778 ATCAGTTTGTCTGAGTCTTTGGTATTTATCTTCTCCATCATGTTTGAAGTGAATG 837
Db 261 IleSerPheValIleuSerAlaIeuValIlePheAsnIeuIeuProIleMetPheGlyValMet 280
Qy 838 CTGTGCACTGGTGTGTTTGTATTACAAATGGGTGCCCATGTTGGTGGTGAACCTTGGGA 897
Db 281 LeuValSerGlyValIleuIyTrIyLysCysGlyAlaGlnPheAlaIeuValIleThrLeuGly 300
Qy 898 ACACCTGGTACATACAGACATTCACAGTTGCAGTCAACCGGTGAGAGATTAATTTAGA 957
Db 301 ThrIeuGlyThrIyTrIyAlaPheThrValAlaValIleThrArgIyThrArgPheArg 320
Qy 958 ATAGAATGAACAAAGCAGATTAATGATGACGTAATGCTGCTATAGACTCACTGCTGAAT 1017
Db 321 IleGluMetAsnIySalIaAspAsnAspAlaGlyAsnAlaIaIleAspSerIeuIeuAsn 340
Qy 1018 TATGAACCTGGAGATTTTAAATATGAAGATATGAAGACAGACATATAGATGATTT 1077
Db 341 TyrIuIyThrValIyTrIyPheAsnAsnGlyuArgIyGlyuAlaGlnArgIyAspGlyPhe 360
Qy 1078 TTGAAGACGTATGAGACTGCTTCACTTGAAGACTCTTACTGCTGATGCTGAACCTT 1137
Db 361 LeuIySerThrIyGluIyThrAlaSerIeuIySerThrSerThrIeuAlaMetIeuAsnIlePhe 380
Qy 1138 GGTCAAAAGTGTATTTTCACTGTCGTTTAAACAGCTAATATGATGTCGCCAGTCAAGGA 1197
Db 381 GlyIleSerAlaIlePheSerValGlyIleuThrAlaIleMetValIeuAlaSerGlnGly 400
Qy 1198 ATTGTGGCAGGTATCCCTTACCTGTGGAGATCTAGTAATGTGAATGATGCTTTTTCAG 1257
Db 401 IleValAlaGlyThrIleuThrValGlyAspIeuValMetValaIleGlyIleuIlePheGln 420
Qy 1258 CTTTCATTATCCCTGAACTTCTGGAACCTATATATAGAGAGATAGCAAGCAAGCTGATA 1317
Db 421 LeuSerIeuProIeuAsnIlePheIleuGlyThrValIyTrArgIuIyTrArgIuIyAlaIleuIle 440
Qy 1318 GATATGAACACTTGTCTTACTCTACTCAAGGTAGACACCAAAATTAAGCAAAAGTATG 1377
Db 441 AspMetAsnThrIleuPheThrIleuIyValaIleAspThrGlnIleIyAspIyValIleThr 460
Qy 1378 GCATCTCCCTTGAGATCACCAACACAGACTACGCTGGCTTGTATATGTGCATTTT 1437
Db 461 AlaSerProIeuGlnIleThrProGlnThrAlaThrValAlaIleAspAsnValaIlePhe 480
Qy 1438 GAATACATTGAGGGCCAGAAAGTCTTAGTGAATATCTTTGAAGTCCCGCAGAGAAAG 1497
Db 481 GluIyTrIleGluGlyGlnIyValaIleuSerGlyIleSerPheGlyuValProAlaGlyLys 500
Qy 1498 AAAGTGGCCATTGTAGAGGTAGTGGGTCAAGGAAAGCAAAATATGAGGCTATTATTT 1557

Db 501 LysValAlaIleValGlyGlySerGlySerGlyLysSerThrIleValaIlePheIleuPhe 520
Qy 1558 CGCTTCTATGAGCCTCAAAAAGGAGTACATTTATCTTCTGGTCAAAATATATCAAGATGTG 1617
Db 521 ArgPheIyGluProGlnIyGlySerIleTyrIleuAlaGlyGlnAsnIleGlnAspVal 540
Qy 1618 AGCTGGAAAACCTTCGAGGGCAGTGGAGTGGTACTTCAGATGCTGCTCTTCAT 1677
Db 541 SerIeuIleuSerIeuAlaGArgAlaValaGlyValValaIleProGlnAspAlaValIlePheHIS 560
Qy 1678 AATTAATTTTATTAACAACCTTTATATGAACAATCAAGTGTCTTCACTGAGGAAGTAT 1737
Db 561 AsnThrIleIyTrIyAsnIleuIeuIyTrGlyAsnIleSerAlaSerProGluIyValIyTr 580
Qy 1738 GCAGTGGCAAAATATAGCTGACCTTCATGATGACATTCCTTGGAAATGCCAATGATATGAC 1797
Db 581 AlaValaIaIySleuAlaGlyIleuHISAspAlaIleIeuIyArgMetProHISGlyTrAsp 600
Qy 1798 ACCCAAGTAGGGGAAACGAGACTCAAGCTTTCAGAGAGAGAAAAGCAAGAGTACAT 1857
Db 601 ThrGlnValaGlyGluArgGlyIleuIySleuSerGlyGlyIyLysGlnArgValaIle 620
Qy 1858 GCAAGACCATTTTGAAGAACCCCAAGTCAATACTTATGATGAAGTACTTACCTGTTA 1917
Db 621 AlaArgAlaIleIeuIySAspProValIleIeuIyTrAspGluAlaThrSerIeu 640
Qy 1918 GATTGCAATCTGGAAGACTATTCCTTGGTGCATGAAGAGATGGTCAACACAGAACT 1977
Db 641 AspSerIleThrGluIyThrIleIeuGlyAlaMetCysAspValaIySHISArgThr 660
Qy 1978 TCTATTTTCATTGACACAGATTTGTCAACAGTGGTGTGATGCAGATGAATCATTTGCTTG 2037
Db 661 SerIlePheIleAlaHISArgIeuSerThrValValaAspAlaAspGluIleIleValIleu 680
Qy 2038 GATCAGGGTAAAGTACCCGAACGTGGTACCACCATGCTTTGCTTGAACCTCATAGT 2097
Db 681 AspGlnGlyIySValaIaGluArgGlyThrHISISeglyIleuIeuAlaAsnProHISer 700
Qy 2098 ATCTATTCAGAAATGTGGCATACACAGACAGCCGTGTGCAAGAACCTGAATACCCCAA 2157
Db 701 IleIySerGluMetCitrPheIyTrGlnSerSerArgValaGlnAsnHISAspAsnProLys 720
Qy 2158 TGGGAA 2163
Db 721 TrpGlu 722
RESULT 5
ID ABB57811
ID ABB57811 standard; Protein; 606 AA.
AC ABB57811;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide seq ID NO 225.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
PN MO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-UTL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX

Qy	2038	GATCAGGGGTAAAGTACCCAGACAGTGGTATCCCACTAGTTGCTTCTACTAACCTCATAGT	2097
Db	556	GIaSGnglYARyValGclYgIunRtgGlyThRHsSerGluLeuLeu---ArgGlnSngly	574
Qy	2098	ATCTATTCGAATAATGGGATACACAGACAGCCGGTGACAGAACCATATTAACCCCAAA	2157
Db	575	LeuTYrAlARgLeuTRpGluThRgInThRgInGlnPhe-----AspProSer	590
Qy	2158	TGGGAAGCAAGAAAGAAATATATCCAAAG	2190
Db	591	ArgGluIleAsnglGluValAlaIaIaIySlys	601
RESULT 6			
ABBS7812	ID	ABBS7812 standard; Protein; 606 AA.	
XX	AC	ABBS7812;	
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 228.	
XX	KW	Drosophila; developmental biology; cell signalling; insecticide;	
XX	KW	pharmaceutical.	
XX	OS	Drosophila melanogaster.	
XX	PN	WO200171042-A2.	
XX	PD	27-SEP-2001.	
XX	PF	23-MAR-2001; 2001WO-US09231.	
XX	PR	23-MAR-2000; 2000US-191637P.	
XX	PR	11-JUL-2000; 2000US-0614150.	
XX	PA	(PEKE) PE CORP NY.	
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
XX	DR	N-PSDB; ABL01915.	
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
XX	PT	interactions -	
XX	PS	Disclosure; SEQ ID NO 228; 21pp + Sequence Listing; English.	
XX	CC	The invention relates to an isolated nucleic acid detection reagent	
XX	CC	capable of detecting 1000 or more genes from Drosophila. The invention	
XX	CC	useful in developmental biology and in elucidating cell signalling and	
XX	CC	cell-cell interactions in higher eukaryotes for the development of	
XX	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
XX	CC	discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA	
XX	CC	sequences (ABLI01840-ABLI6175) and the encoded proteins	
XX	CC	(ABBS7737-ABBS72072).	
XX	CC	The sequence data for this patent did not form part of the printed	
XX	CC	specification, but was obtained in electronic format directly from WIPO	
XX	CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX	XX	Sequence 606 AA;	
Alignment Scores:			
	Pred. NO.:	7.51e-165	Length: 606
	Score:	1852.00	Matches: 360
	Percent Similarity:	76.92%	Conservative: 110
	Best Local Similarity:	58.92%	Mismatches: 131
	Query Match:	43.54%	Indels: 10
	DB:	22	Gaps: 3
AF133659 (1-2345) x ABBS7812 (1-606)			

QY	358	ATCATAAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGGCCAGACTTACGAGCTAGA	417
QY	358	ATCATAAAGCAATGCTTTCTTATGTGTGGCCCAAGACAGGCCAGACTTACGAGCTAGA	417
Db	1	MetLeuAaGlaMetMetAlaTyrIleTrpProlyGluAspProLeuValAArglySArg	20
QY	418	GTTCGCAATTCGGCTGGGATTTTGGGTGGTGCAGAAAGCCATGAAATATGTGGTTCCTTC	477
Db	21	ValGlyIleSerLeuGlyLeuLeuAlaGlySerLysLeuLeuThrValCysValProPhe	40
QY	478	ATGTTTAAATGTGCTGTAGACAGCGCTCAACCAAGATGCGGAAACATGCTGAACCTGAGT	537
Db	41	LeuPheLysGlyAlaValAspThrMet-----ThrThrLeuAsnMetAsp	55
QY	538	GATGCACCAAAATACAGTTTGCAACCATGGACAGACAGCTTCTGATGGCTATGATCA	597
Db	56	ThrAlaProAspAlaValIleuSerAlaIaThrAlaLeuMetLeuGlyTyrGlyIleAla	75
QY	598	AGAGCTGGAGCTGCTTTTTTTAAAGAAAGTTCCAAATGCGAGTATTGGCAAGTAGCCGACG	657
Db	76	ArgAlaSerAlaIaAlaGlyPheAsnGluLeuValArgAsnAlaValPheAlaLysValAlaHis	95
QY	658	AATTCATCCCAAGAAATAGCCCAAAATATGTCTTCTCCATCTTCACACACCTGGATCGGGT	717
Db	96	HisSerIleArgLysIleAlaSerAsnValAPheLeuHisIleuHisAsnLeuSpleuAla	115
QY	718	TTTCACTGACAGACAGACAGCGGAGCTTTATCTTAAGGCTATTGACAGAGAAACAAGGGT	777
Db	116	PheHisLeuAsnLysGlnIntrGlyAlaLeuSerLysThrIleAspArgGlySerArgGly	135
QY	778	ATCAGTTTGTCTGAGTGCCTTTGGTATTATCTTCTTCCATCATGTGTTGAAGTATG	837
Db	136	IleAsnPheValIleuSerAlaMetValPheAsnIleValProThrIlePheGluLeuAla	155
QY	838	CTTGCACTGGTGTTTGTATTACAATAGCGGGGCCAGCTTTGGCTTGGTAACCCCTGGA	897
Db	156	LeuValSerSerIleLeuGlyValLysCysGlyLeuAlaPheAlaValSerMetGly	175
QY	898	ACACTTGGTACATTCACAGCATTCACAGTTGCAGTCACACGGTGGAGAACTGATTTAGA	957
Db	176	CysValGlyIleTyrAlaIaLysTyrThrLeuSerValThrGlnTrpArgThrGlnPheArg	195
QY	958	ATGAAATGAAACAAGACAGATTAATGATGACGAGTAATGCTGATTAAGACTCAGCTGCAAT	1017
Db	196	ValPheMetAsnGlnIaGluAsnGlnIaGlyAsnLysAlaValAspSerLeuIleAsn	215
QY	1018	TATGAACCTGGGAAGTATTTTAATATGAAGAATATGAAGCACAGAGATATGATGATT	1077
Db	216	TyrGluThrValLysTyrPheAsnAsnGluLysTyrGluAlaGlyCysTyrAsnGluVal	235
QY	1078	TTGAAAGACGTATGAGACTGCTGCTTATGAAGAAAGTACTACTGCTGCTATGCTGAACCTT	1137
Db	236	LeuLysLysTyrGluAlaIaIaSerLeuLysThrSerSerLeuAlaIleuLeuAsnPhe	255
QY	1138	GGTCAAAAGTGTCTATTTTCAAGTGTGCGTTTAAACAGTATAATGAGTGTGCCAGTCAGGA	1197
Db	256	GlyIleAsnAlaIaIePheSerSerAlaLeuSerLeuIleMetValLeuAlaIaLysGlu	275
QY	1198	ATTGTGGCAGAGTACCCCTTACTGTTGGAGATCTAAGTAATGGTGAATGACTGCTTTTCAG	1257
Db	276	IleIaIaGlnGlyAsnMetThrValAGlyAspLeuValMetValAsnAlaIleuLeuPheGln	295
QY	1258	CTTTCATTAACCCCGGAACCTTCTGGGAAACGTATATAGAGAGACATACAGACACTCATTA	1317
Db	296	LeuSerIleProLeuGlyPheLeuGlySerValTyrArgGluValAArgGlnAlaLeuLeu	315
QY	1318	GATATGAACACTTGTGTTACTCTTACTTCAAGAGTATAGACACCCCAATTTAAGCAAAAGTGATG	1377
Db	316	AspMetArgAlaMetPheThrIleuMetAsnValAspSerSerIleGlnThrAlaIaIaAsn	335
QY	1378	GCATCTCCCTTCAGATCACACACAGACAGACTACCGTGCCCTTGGATATGTGCATTTT	1437
Db	336	AlaGlnProLeuPheValAspThrThrAsnSerSerIleGluPheArgAsnValSerPhe	355

```

QY 1438 GAATACATTAGAGGCGCAAGAAAGTCTTAGTGAATATCTTTGAAGTCCCTGACAGAAAG 1497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 356 GUTYRGUProGlyLysProIlePhehArgAspLeuSerPheThrIleProIleGlyLys 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1498 AAAGTGGCCATTGTAGAGGTAGGCTGAGGAAAGCAATATGTAGAGCTATTATT 1557
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 376 AsnValAlaIleValAlGlyLysSerGlyLysSerSerMetValArgLeuLeuPhe 395
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1558 CGCTCTTAGAGCTTCAAAAGGTAGCATTTTCTTGCTGCTCAAAATATACAGATGG 1617
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 396 ArgPhePheGluProAsnSerGlyLysValLeuIleGlyGlyIleAsnAlaVal 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1618 AGCTTGAAAGCCTTGCAGAGGACGTGAGTGTACTCAGAGTCTGCTCTTCAT 1677
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 416 AspLeuGluSerLeuArgLysValIleAlaValProGlnAspSerValLeuPheHis 435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1678 AATCTATTATTACAACTCTTATATGAAACATCATGCTTCACTGAGAGAGTAT 1737
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 AsnThrIleGluHisAsnIleHisIleSTYrgLysAsnLeuSerLysSerHisAlaGluValGln 455
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1738 GCAAGTGGCAAAATAGCTGCAATGATGATGCAATTTCTGCAATGCAATGATATGAC 1797
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 456 AsnAlaAlaArgMetAlaAspLeuHisAspSerIleMetSerTrpProGlyGlnTyrSer 475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1798 ACCCAAGTAGGGGAAAGAGAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 476 ThrGlnValGlyGlyAlaGlyLeuLysLeuSerGlyGlyLysGlnAlaArgValAlaIle 495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1858 GCAAGAGCCATTGTGAAGACCCCGAGTCACTATCTATGATGAAAGCTTCTCATCTGTTA 1917
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 496 AlaArgAlaIleLeuLysAsnThrProIleLeuIlePheAspGluAlaThrSerSerLeu 515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1918 GATTGATTACTGGAAGAGACTTATCTTGCTGATGAGAGAGAGAGAGAGAGAGAG 1977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 516 AspSerIleThrGlnHisAsnIleLeuGlnAlaLeuThrArgAlaThrSerGlyArgThr 535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1978 TCTATTTCATTGCACACAGATTTGTCAACAGTGTGATGAGAGAGAGAGAGAGAG 2037
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 536 SerIleCysIleAlaHisIleArgLeuSerThrValLysAspAlaAspIleLeuValLeu 555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2038 GATCAGGGTAGAGTAGCCGAGCTGTGATCCACCATGTTGCTTCTGCTTACCTCATAGT 2097
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 556 GUAANGlyArgValAlGlyGlyLysArgGlyThrHisSerGluLeuLeu---ArgGlnAsnGly 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2098 ACTCTATTCAGAAATGTGGATACACAGACAGCCGCTGTGCAGAACCATATTAACCCCAA 2157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 575 LeuTyrAlaArgLeuTrpGluThrGlnTrpGlnPhe-----AspProSer 590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2158 TGGGAAGCAAGAAAGAAATATATATCCAAAGAG 2190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 591 ArgGluIleAsnGluGluValAlaAlaLysLys 601
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

PF 05-MAR-2001; 2001WO-EP02478.
XX
PR
XX 06-MAR-2000; 2000EP-0104770.
XX
PA (MERE ) MERCK PATENT GMBH.
XX
PI Brandt S;
XX
DR WPI: 2001-589940/66.
DR N-PSDB: AAS15604.
XX
PT Novel ABC transporter polypeptides for the treatment of cystic
XX fibrosis, adrenoleukodystrophy, hyperglycaemia and hypertension
XX
PS Claim 1; Page 38-39; 44pp; English.
XX
CC This sequence represents the human ABC transporter expressed in liver
CC (ATIL) protein of the invention. The invention also comprises the
CC sequence of the ATIL cDNA, an agonist or antagonist of this protein, an
CC expression system which can be used to express recombinant ATIL protein
CC or a fusion protein comprising an immunoglobulin Fc-region in a host
CC cell and an antibody immunospecific for the ATIL protein which can be
CC used to stimulate or inhibit a function. This protein may have
CC hypotensive activity and the sequences can be used in gene therapy or to
CC develop a vaccine to treat diseases, including conditions where cellular
CC resistance to toxins is induced, drug efflux is altered, intracellular
CC distribution and antigen presentation pathways are involved, such as
CC disorders related to lipid metabolism and heavy metal transport, cystic
CC fibrosis, adrenoleukodystrophy, hyperinsulinaemic hypoglycaemia and
CC hypertension. The nucleotide sequence is useful as a diagnostic reagent
CC for detecting mutations in associated genes, in chromosome localisation
CC studies, and as a valuable tool for tissue expression studies. The
CC sequences may also be useful as vaccines, and for inducing an
CC immunological responses in a mammal and are useful to configure
CC screening methods for detecting the effect of added compounds on the
CC production of mRNA and polypeptide in cells. The ATIL protein can also
CC be useful to identify membrane-bound soluble receptors.
XX
SQ Sequence 842 AA;

```

Alignment Scores:

```

Pred. No.: 2.95e-107 Length: 842
Score: 1239.50 Matches: 286
Percent Similarity: 53.86% Conservative: 126
Best Local Similarity: 37.39% Mismatches: 240
Query Match: 29.14% Indels: 113
DB: 22 Gaps: 16

```

AF133659 (1-2345) x AAU09965 (1-842)

```

QY 10 CTCGGGATGCAATTTCTGCGCTGGGCGGCGGCGCTTTCGAAAGCCGCCGAC 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 136 LeuAlaMetGlyIleTrp-----IleLysPheArgHis 146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 70 TCC-----GCGATTCTGATCCGGCTTCTTACTCTGTT 102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 147 SerProGlyLeuLeuLeuLeuTrpThrValAlaPheAlaIleGluAsnLeuAlaLeuVal 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 103 AGCGCTCAGGTCGCGCATGAGAGGCCACATCACTCGCGCTTGGAAACCGCTCGAGCC 162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 167 SerTrpAsnSerProGlnTrp----- 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 TACCGATTCGAGAGTCAATTAAGATATCAATGCGAGAGATTGGAAAGCAATTC 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 174 -----TrpTrpAlaArgAlaAspLeuGlyGlnGln 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 223 GCAGAGTCTTATGATGCTGCAAAAGGCTCTCCAGATATGCGCATGATGAAAGAGACA 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 184 ValGlnPhe-----SerLeuTrpValLeu----- 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 283 TGTGGCATGTGATGACGAGAGAGAGACTCCACACA----- 318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 192 -----ArgTyrValValSerGlyGlyLeuPheValLeuGlyLeuTrpAlaProGlyLeu 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 319 -----GACCCAAAGAGGTTAAAAAGATTGTAATCTCGGAAA 357
Db 210 ArgProGlnSerIYrThrLeuGlnValHlsglGlnIuaSerGlnAspValGlnArgSerGln 229
QY 358 ATC-----ATMAAGCATG 372
Db 230 ValArgSerAlaAlaGlnGlnSerThrTrpArgPheGlnArgValArgValLeuArgLeuLeu 249
QY 373 CTTTCTTATGTGTGGCCCAAGACAGCCAGATCTACAGCTAGAGTTGCCATTTGCTG 432
Db 250 SerGlyrYrLeuTrpProArgGlySerProAlaLeuGlnLeuValValLeuIleCysLeu 269
QY 433 CGATTTTGGGTGTGCAAGAGCCATGAATATTGTTCCCTTCAGTTTAAATATGCT 492
Db 270 GlyLeuMetGlyLeuGlnArgAlaLeuAsnValLeuValProIlePheTyrArg----- 287
QY 493 GTACAGACCCCAACCAAGATGTCGGAAACATGCTGAACCTG-----AGTGAAGCACCA 546
Db 288 -----AsnIleValaLeuLeuLeuThrGlnIuLysAlaPro 298
QY 547 ---AATACAGTTGCACCATGCAACAGCA-----GTTCTGATGGCTAT 588
Db 299 TrpAsnSerLeuAlaIleTrpThrValIleSerTyrValPheLeuLysPheLeuGlnGly 318
QY 589 GGTGTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTGCAATGCAATTTGGCAG 648
Db 319 GlyThrGlySerThrGly-----PheValSerAsnLeuArgThrPheLeuTrpIleArg 336
QY 649 GTACCCAGAAATTCATCCCAAGAAATAGCCAAAATGCTTTCCTTCATCTTCACAACTG 708
Db 337 ValGlnGlnPheThrSerArgArgValGlnLeuLeuIlePheSerHlsLeuHlsGlnLeu 356
QY 709 GATCTGGGTTTCACTGAGCAGACAGACGGAGCTTATCTTAAGCTATGACAGAGA 768
Db 357 SerLeuArgTrpHlsLeuGlyArgArgThrGlyGlnValLeuAlaGlnIleAlaAspArgGly 376
QY 769 ACAAGGGGATCACTTTTCTGAGTGCCTTTGATTTAACTCTTCCCATGCTTT 828
Db 377 ThrSerSerValThrGlyLeuLeuSerTyrLeuValPheAsnValIleProThrLeuAla 396
QY 829 GAATGATGCTGTGCAGTGGCTTTTGATTAACAATGC-----GGTCCCACTTTGCT 882
Db 397 AspIleIleIleIle-----GlyIleIleIleTyrPheSerMetPhePheAsnAlaIleTrpPheGly 414
QY 883 TTGTAACCTTGGAAACCTGTATACACAGCATTCACAGTTGCAGTGCACGCGTG 942
Db 415 LeuIleValPheLeuCysMetSerLeuTyrLeuThrLeuThrIleValIleValThrGlnTrp 434
QY 943 AGAATGATTTAGAAATGAATGAACAAAGCAGATATGACAGTAATGCTGTATA 1002
Db 435 ArgThrLysPheArgGlnAlaMetAsnThrGlnGlnAsnAlaThrArgAlaArgAlaVal 454
QY 1003 GACTCATCTGTAATTTGAACCTGTGAAGTATTATTAATGAAGATATGAAGACACAG 1062
Db 455 AspSerLeuLeuAsnPheGlnThrValIleTyrTyrAsnAlaGlnSerTyrGlnIuValGln 474
QY 1063 AGATATGATGATTTTGAAGAGATGAGACCTGCTTCAATGAAGATACCTTAACCTG 1122
Db 475 ArgTyrArgGlnAlaIleIleIleLysTyrGlnGlyLeuGlnTrpLysSerSerAlaSerLeu 494
QY 1123 GCTATGCTGAACCTTTGGTGAAGTGTATTTTCAGTTCGCTTAAACAGATTAATGCTG 1182
Db 495 ValLeuLeuAsnGlnThrGlnAsnLeuValIleGlyLeuGlyLeuLeuAlaGlySerLeu 514
QY 1183 CTGCCAGCTCAGGGAATTTGGCAGGTACCTTACTGTTGAGATCTAGTATGCTGAAT 1242
Db 515 LeuCysAlaTyrPheValIleThrGlnGlnLysLeuGlnValGlyAspTyrValLeuPheGly 534
QY 1243 GGAATGCTTTTTCGCTTATTAACCCCTGAACCTTCTGGGAATGATATATGAGAGACT 1302
Db 535 ThrTyrIleIleGlnLeuTyrMetProLeuAsnTrpPheGlyTrpTyrArgMetIle 554

```

```

QY 1303 AGACACACATCTATGATATGAACACACTTGTATTACTTACATCAAGGTAGACACCCAAATT 1362
Db 555 GlnThrAsnPheIleAspMetGlnAsnMetPheAsnLeuLysGlnGlnThrGlnIuVal 574
QY 1363 AAAGACAAAGATGATGATCTCCCTCAGATCAACACACAGACGATACCGTGCCCTT 1422
Db 575 LysAsnLeuProGlyAlaGlyProLeuArgPhe-----GlnLysGlyArgIleGlnPhe 592
QY 1423 GATATGTCATTTTGAATCATTTAGAGGCGCAGAAAGCTCTTATGGAATATCCTTTGAA 1482
Db 593 GlnAsnValHlsPheSerTyrAlaAspGlyArgGlnThrLeuGlnAspValSerPheThr 612
QY 1483 GTCCCTCAGAGAAAGAAAGTGGCCATTGTAGAGAGTGTGGCTCAGGAAAGACAAATA 1542
Db 613 ValMetProGlyGlnThrLeuAlaLeuValGlyProSerGlyAlaGlyLysSerThrIle 632
QY 1543 GTGAGCTATTTATTTGCTTCTATGACCTCAAAAGGTAGCAATTATCTTGCTGTCAA 1602
Db 633 LeuArgLeuLeuPheArgPheTyrAspIleSerSerGlyCysIleArgIleAspGlyGln 652
QY 1603 AATATCAAGATGTGAGCCCTGGAAGACCTTCGAGGCGACGTGGAGTGTACCTCAGAGT 1662
Db 653 AspIleSerGlnValThrGlnAlaSerLeuArgSerHlsIleGlyValIleProGlnAsp 672
QY 1663 GCTGCTCTTCCATATATCTATTATTATTAACACCTCTTATATGGAACATCAGGCTTCA 1722
Db 673 ThrValLeuPheAsnAspThrIleAlaAspAsnIleArgTyrGlyArgValThrAlaGly 692
QY 1723 CTTGAGAAATGTATGTAGAGTGGCAAAATTAGCTGCACTTCAATGATGCAATTCTTCAATG 1782
Db 693 AsnAspGlnValGlnAlaAlaGlnAlaAlaGlyIleHlsAspAlaIleMetAlaPhe 712
QY 1783 CCACATGATATGACACCCCAAGTAGGGGAACGAGACTCAAGCTTTCAGAGAGAGAAAG 1842
Db 713 ProGlnGlyrYrArgThrGlnValGlyGlnArgGlyLeuLysLeuSerGlyGlyGlnLys 732
QY 1843 CAAGAGTACCAATTGCAAGAGCCATTTTGAAGACCCCTCAGTATCTATGATGAA 1902
Db 733 GlnArgValAlaIleAlaArgThrIleLeuLysAlaProGlyIleIleLeuLeuAspGln 752
QY 1903 GCTAATTCATCGTTAATTTGATTTACTGAGAGACATTTCTTGCGCATGAAAGATGTG 1962
Db 753 AlaThrSerAlaLeuAsnThrSerAsnGlnArgAlaIleGlnAlaSerLeuAlaLysVal 772
QY 1963 GTCAAACACAGCTTCTATTTTCATTTGACACAGATTGCAACAGTGTGATGAGAT 2022
Db 773 CysAlaAsnArgThrThrIleValIleAlaHlsArgLeuSerThrValIleAsnAlaAsp 792
QY 2023 GAATCATTTGCTTGGATCAGGTAAAGGTAGCCGAACGTGTACCCACATGTTGCTT 2082
Db 793 GlnIleLeuValIleLysAspGlyCysIleValGlnArgGlyArgHlsGlnAlaLeuLeu 812
QY 2083 GCTAACCTCATATGATCTATTCAAAATTTGGCATATACACAGACAGCCGTGCGAGAAC 2142
Db 813 SerArg---GlyGlyValTyrAlaAspMetTrpGlnLeuGlnGlnGlnGlnGlnThr 831
QY 2143 CATGATACCCCAAA 2157
Db 832 SerGlnAspThrLys 836

```

RESULT 8
 AAE16764
 ID AAE16764 standard; protein; 842 AA.
 XX
 AC AAE16764;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human transporter and ion channel-1 (TRICH-1) protein.
 KW Human; transporter and ion channel-1; TRICH-1; neuroprotective; asthma;
 KW nootropic; cytosolic; cardiovascular; immunosuppressive; cardiomyopathy;
 KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;

KM Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;
 KM amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;
 KM Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;
 KM cell proliferative disorder; psoriasis; cardiac disease; hypertension;
 KM bradyarrhythmia; gene expression; drug screening.

XX Homo sapiens.

XX Key location/Qualifiers

XX Domain 186..203

XX Domain /label= Transmembrane_domain

XX Domain 386..406

XX Domain /label= Transmembrane_domain

XX WO200192304-A2.

XX 06-DEC-2001.

XX 25-MAY-2001; 2001WO-US17065.

XX 26-MAY-2000; 2000US-208424P.

XX 01-JUN-2000; 2000US-209001P.

XX 08-JUN-2000; 2000US-210588P.

XX 16-JUN-2000; 2000US-212335P.

XX 22-JUN-2000; 2000US-213747P.

XX 29-JUN-2000; 2000US-215391P.

XX (INCYTE GENOMICS INC.

XX Thornton M, Walla NK, Yue H, Nguyen DB, Lai P, Gandhi AR;

XX Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YN;

XX Azimzal Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS;

XX Rammann BE, Lee EA, Hafalia A, Baughn MR, Green BD, Khan PA;

XX Kearney L, Elliott VS, Sellhammer JJ, Policky JL, Borowsky ML;

XX Burford N, Ding L, Lu DM, Hillman JL;

XX WPI; 2002-122055/16.

XX N-PSDB; AAD27254.

XX New human transporters and ion channels (TRICH) polypeptides useful for

XX diagnosing, treating or preventing disorders associated with aberrant

XX expression of TRICH

XX Claim 1; Page 144-146; 210pp; English.

XX The invention relates to human transporters and ion channels (TRICH)

XX polypeptides and their cDNA molecules. The nucleic acid and polypeptide

XX sequences are useful in the diagnosis, treatment, and prevention of

XX disorders associated with transport (akinesia, cystic fibrosis, Bell's

XX palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,

XX amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's

XX muscular dystrophy); immunological (AIDS, Addison's disease, allergies,

XX asthma); cell proliferative disorders (cancers, leukemia, psoriasis);

XX cardiac disease (angina, hypertension, or bradyarrhythmia) and in the

XX assessment of the effects of exogenous compounds on the expression of

XX nucleic acid and amino acid sequences of transporters and ion channels.

XX The polynucleotides may be used to detect and quantify gene expression

XX in biopsied tissues in which TRICH expression may be correlated with a

XX disease, to generate hybridization probes for mapping naturally occurring

XX genomic sequence, and in drug screening. The present sequence is human

XX TRICH-1 protein.

XX Sequence 842 AA;

XX Alignment Scores:

XX Pred. No.: 2,956-107

XX Score: 1239..50

XX Percent Similarity: 53.86%

XX Best Local Similarity: 37.39%

XX Query Match: 29.14%

XX DB: 23

XX Length: 842

XX Matches: 286

XX Conservative: 126

XX Mismatches: 240

XX Indels: 113

XX Gaps: 16

AF133659 (1-2345) x AAE16764 (1-842)

QY 10 CTCGCGATGCAATTCCTGGCGCTGGCGCGCGCGCTGCTTTCGAAAAGCCGCGCAC 69
 |||||
 DB 136 LeuAlaMetGlyIleTrp-----IleYsPheArgHis 146
 QY 70 TCC-----CGAATTCGATCCGCGCTTATGCTTCTT 102
 |||||
 DB 147 SerProGlyLeuLeuLeuLeuTrpThrValAlaPheAlaIaGluAsnLeuAlaLeuVal 166
 QY 103 AGCGGCTCAGGTCGCGAGTGGAGGCGACATCACTCGGCGCTTGGGAACCGCTCGAGCC 162
 |||||
 DB 167 SerTrpAsnSerProGlnTrp----- 173
 QY 163 TACCAAGATTCACAGATCAATTAAGATATCATGCGACAGATTGGGAAAAGCAATTCA 222
 |||||
 DB 174 -----TrpThrPAlaArgAlaAspLeuGlyGlnGln 183
 QY 223 GCACAGTCTTATGATGCTGCAAGAGCTCTCCAGGTATGCGCATGATGAAAAGAGACA 282
 |||||
 DB 184 ValGlnPhe-----SerLeuTrpValLeu----- 191
 QY 283 TGTGGCATGTCATGACGAGAGAGACTCCACACA----- 318
 |||||
 DB 192 -----ArgTrpValValSerGlyGlyLeuPheValLeuGlyLeuTrpAlaProGlyLeu 209
 QY 319 -----GACCCAAAAGAGGGTTAAAGATGTTGATATCTCGAAA 357
 |||||
 DB 210 ArgProGlnSerTrpThrLeuGlnValHisGlnGluAspGlnAspValGluArgSerGln 229
 QY 358 ATC-----ATAAAGCAATG 372
 |||||
 DB 220 ValArgSerAlaIaGlnGlnSerThrTrpArgAspPheGlyArgLeuValLeuValLeuLeu 249
 QY 373 CTTTCTTATGTCGGCCCAAGACAGCGACATATGACAGTATGATGTCATGTCGCTG 432
 |||||
 DB 250 SerGlyTrpLeuTrpProArgGlySerProAlaLeuGlnLeuValValLeuIleCysLeu 269
 QY 433 GGATTTTGGGTGGTGCAAGAGCCATGATATTTGTTGCTTCATGTTAAATATGCT 492
 |||||
 DB 270 GlyLeuMetGlyLeuGluArgAlaLeuAsnValLeuValProIlePheTrpArg----- 287
 QY 493 GTAGACAGCTCAACCAAGATGCGGAAACATGCTGAACCTG-----AGTAGTCACCA 546
 |||||
 DB 288 -----AsnIleValAsnLeuLeuThrGlnIuValAlaPro 298
 QY 547 ---AATACAGTTCACACCATGCGCAACACA-----GTTTCGATGGCGTAT 588
 |||||
 DB 299 TrpAsnSerLeuAlaTrpThrValThrSerTrpValPheLeuLysPheLeuGlnGlyGly 318
 QY 589 GGTGATATCAAGAGCTGAGCTGCTTTTAAACGAAGTTCGAAATGCAATTTGGCAAG 648
 |||||
 DB 319 GlyThrGlySerThrGly-----PheValSerAsnLeuAlaGlyThrPheLeuTrpIleArg 336
 QY 649 GTAGCCCAAGATTCATATCGAAGAAATAGCCAAAATGCTTTTCATCTTCACACACTG 708
 |||||
 DB 337 ValGlnGlnPheThrSerArgValGluLeuLeuIlePheSerHisLeuHisGlyLeu 356
 QY 709 GATCTGGGTTTTCACCTGACAGACAGACCGGACCTTATATTAAGCGTATTGACAGAGA 768
 |||||
 DB 357 SerLeuAlaGlyTrpHisLeuGlyAlaArgThrGlyIuValLeuAspGlyIleAlaAspArgGly 376
 QY 769 ACAAGGGGTATCAGTTTGTCTGAGTCTTGTGATTTATCTTCCTCCATCAATGTT 828
 |||||
 DB 377 ThrSerSerValThrGlyLeuLeuSerTrpLeuValPheAsnValIleProThrLeuAla 396
 QY 829 GAAGTATGCTTGTACAGTGTGTGTTGTATTAACAATGC-----GGTCCCAAGTTGCT 882
 |||||
 DB 397 AspIleIleIle-----GlyIleIleTrpPheSerMetPhePheAsnAlaIleTrpPheGly 414
 QY 883 TTGGTAACTTCTTGAACACCTGTGATACACACAGATTCACAGTTGCGATACACGCGTG 942
 |||||
 DB 415 LeuIleValPheLeuCysMetSerLeuTrpLeuThrIleValIleValIleValThrGluTrp 434

The sequence represents a human ATP binding cassette (ABC) transporter family member, ABCB12. The nucleic acid encoding ABCB12 is useful as a target for developing modulating agents of multidrug resistance and as diagnostic and therapeutic tools to control cellular transport in cells of the brain and transport across the blood-brain-barrier and for developing strategies to deliver drugs to the brain, for treating Alzheimer's disease, hyperinsulinaemic hypoglycaemia, cystic fibrosis, atherosclerosis, macular degeneration, adrenoleukodystrophy, prion diseases, Huntington's disease, Parkinson's disease and for treating mood and panic disorders. The nucleic acid is useful for expressing ABCB12, to detect ABCB12 transporter mRNA or genetic alteration in an ABCB12 transporter gene, and to modulate ABCB12 transporter activity and is useful in tissue-typing to identify individuals in forensic biology. The protein is useful as an immunogen to raise anti-ABCB12 transporter antibodies and is used to screen for naturally occurring ABCB12 transporter substrates and to screen for drugs or compounds which modulate ABCB12 transporter activity. The nucleic acid and protein are used to treat a subject having a disorder characterised by aberrant or unwanted ABCB12 transporter protein or nucleic acid expression or activity. The nucleic acid, protein and antibodies against ABCB12 are useful in screening assays, predictive medicine and in prophylactic and therapeutic treatment. The antibodies are useful to detect and isolate ABCB12, regulate the bioavailability of ABCB12 and to modulate ABCB12 transporter activity. A vector expressing the nucleic acid is useful for producing non-human transgenic animals.

XX Sequence 843 AA:

Alignment Scores:

Pred. No.:	1,85e-106	Length:	843
Score:	1231.00	Matches:	284
Percent Similarity:	53.57%	Conservative:	129
Best Local Similarity:	36.84%	Mismatches:	250
Query Match:	28.94%	Indels:	108
DB:	22	Gaps:	16

AF133659 (1-2345) x AAU00010 (1-843)

```

QY 10 CTGCGAGTCATCTTCTGGCGCTGCGCGCGCGCTTTCGAAAGCGCGCAC 69
DB 136 LeuAlaMetGlyIleTrp-----IleYsPheAlaGHis 146
QY 70 TCC-----GGCATTCGATCCGCGCTTACGCTCTCTGT 102
DB 147 SerProGlyLeuLeuLeuLeuTrpThrValAlaPheAlaIleGluAsnLeuAlaLeuVal 166
QY 103 AGCGGCTCAGTCCGCGATGAGGCCACATCCTCGCGCTTGGAACCGCTCGAGCC 162
DB 167 SerTPAsnSerProGlnTrp----- 173
QY 163 TACGAGATTCGAGATCATTAAAGTATCACATGCGACAGATTGGGAAGCAATTCA 222
DB 174 -----TTPTPAlaArgAlaAspLeuGlnGln 183
QY 223 GGACAGTTCTTAGAGTCGCAAAAGGCTCTCCAGGTATGCGCCATGATAAAAGAGACA 282
DB 184 ValGlnPhe-----SerLeuTrpValLeu----- 191
QY 283 TGTGGCATGTCATGACGAGAGAGAGACTCCACACA----- 318
DB 192 -----ArgTyrValValSerGlyLeuPheValLeuGlyLeuTrpAlaProGlyLeu 209
QY 319 -----GACCCAAAGAAGGTTAAAGATCTTATACTCGGAAA 357
DB 210 ArgProGlnSerTyrThrLeuGlnValHisGlnGluAspValAlaGluArgSerGln 229
QY 358 ATC-----ATTAAGCAATG 372
DB 230 ValArgSerAlaAlaGlnInSerThrTPArgAspPheGlyArgIleYsLeuArgLeuLeu 249
QY 373 CTTTCTTATGTCGCGCAAGACAGCCAGATCTACGAGCTAGAGTTGCCATTTCGCTG 432
DB 250 SerGlyTyrLeuTrpProArgGlySerProAlaLeuGlnLeuValValLeuIleCysLeu 269

```

```

QY 433 GAATTTTGGGNGCAAGGCCATGATATGTTGCTTCATGCTTAAATATGCT 492
DB 270 GlyLeuMetGlyLeuGluArgAlaLeuAsnValLeuValProIlePheTyrArgAspIle 289
QY 493 GTAGACACCTCAACAGATCTCGGAAAACATGCTGAACCTGAGTATGACCA--AAT 549
DB 290 ValAsnLeuLeuThrGlu-----LysAlaProTrpAsn 300
QY 550 ACAGTTGCA---ACATGGCAACAGACGTTCTGATGCTAT-----GGGTATCAAGA 600
DB 301 SerLeuAlaTrpThrValThrThrTyrValaPheLeuYsPheLeuGlnGlyGlyThr 320
QY 601 GCTGAGCTGCTGTTTTTAAACGAATTGCAATGACGATTTGGCAAGTGGCCAGAAAT 660
DB 321 GlySerThrGlyPheValSerAsnLeuArgThrPheLeuTrpIleArgValGlnGlnPhe 340
QY 661 TCATTCGAGAAATATGACCAAAATGTCTTCTCCATCTTCACAACTGGATCTGGGTTT 720
DB 341 ThrSerArgArgValGluLeuLeuIlePheSerHisLeuHisGluLeuSerLeuArgTrp 360
QY 721 CACCTGACACAGACACGCGAGCTTTATCTTAAGCTATTCACAGAGAACAAAGGGTATC 780
DB 361 HisLeuGlyArgArgThrGlyGluValLeuArgIleAlaAspArgGlyThrSerSerVal 380
QY 781 AGTTTGTCTGAGTGTGTTGATTAATCTCTCCATCATGCTTGAAGTATGCTT 840
DB 381 ThrGlyLeuLeuSerTyrLeuValaPheAsnValIleProThrLeuAlaAspIleIle 400
QY 841 GTCACTGTGTTTGTATTAACAATGC-----GGTCCACGTTTGTCTTGTAACCTT 894
DB 401 -----GlyIleIleTyrPheSerMetPhePheAsnAlaTrpPheGlyLeuIleValaPhe 418
QY 895 GGAACACTGTGTACATACACAGACTTCACACTTCAGCTGACACGCTGAGAACTAGATT 954
DB 419 LeuCysMetSerLeuTyrLeuThrLeuThrIleValaValThrGluTrpArgThrIleYsPhe 438
QY 955 AGAATGAATAATGACCAAGCAAGCATATGATGACGATATGCTGATAGACTCATGCTG 1014
DB 439 ArgArgAlaMetAsnThrGlnGluAsnAlaThrArgAlaArgAlaValaAspSerLeuLeu 458
QY 1015 AATTATGAACCTGTGAAGTATTTAATATGAAGATATGAAGCACACAGATATATGATGA 1074
DB 459 AsnPheGluThrValValTyrTyrAsnAlaGluSerTyrValGluValArgTyrArgGlu 478
QY 1075 TTTTGAAGAGTATGACACTGCTTATGAAGAACTACCTCTACCTGCTATGCTGAAAC 1134
DB 479 AlaIleIleTyrTyrGlnGlyLeuGluTrpIleYsSerSerAlaSerLeuValLeuLeuAsn 498
QY 1135 TTGTGTCAAAGTCATATTTTCAGTGCCTTTTAAACAGCTATTAATGCTGCTGCCAGTCA 1194
DB 499 GlnThrGlnAsnLeuValIleGlyLeuGlnIleuValaGlyAspTyrValLeuPheGlyThrTyrIleIle 518
QY 1195 GGAATTTGGCAGGTATCCCTTACTGTTGAGATCTAGTAATGTGTAATGACTGCTTTT 1254
DB 519 PheValThrGlnGlnIleuLeuGlnValaGlyAspTyrValLeuPheGlyThrTyrIleIle 538
QY 1255 CAGCTTCAATTAACCTGTAACCTTCTGGAACTGTATATAGAGACTAGACAAAGACATC 1314
DB 539 GlnLeuTyrMetProLeuAsnTrpPheGlyThrTyrTyrArgMetIleGlnThrAsnPhe 558
QY 1315 ATAGATATGAACACCTGTTTACTACTACTCAAGGATGACCCCAATTAAGACAAAGTG 1374
DB 559 IleAspMetGluAsnMetPheAspLeuLeuYsGluGlnThrGluValaLysAspLeuPro 578
QY 1375 ATGCGATCTCCCTTCAGATCAACACAGACAGCTACCGTGGCTTTGATATATGTGAT 1434
DB 579 GlyAlaGlyProLeuArgPhe-----GlnIlysgIArgIleGluPheGluAsnValHis 596
QY 1435 TTTGAATACATTAAGAGGCCAAGAAAGTCTTATGATGATATCTTTGAAGTCCCTGCACGA 1494
DB 597 PheSerTyrAlaAspGlyArgGluThrLeuGlnAspValSerPheThrValMetProGly 616

```



```

Qy 1495 AAGAAAGTCCATTGTGAGAGTGGTCCAGGAGAAACACAAATGATGAGCCTATTA 1554
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 617 GlnThrIleuAlaIeuValGlyProSerGlyAlaGlySerThrIleLeuArgLeu 636
Qy 1555 TTGGCTTCCTAGAGCCCAAAAGGTAGCATTTATCTGTGCTGCAAAATATCAAGAT 1614
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 637 PheArgPheIyrAspIleSerSerGlyCysIleArgIleAspGlyGlnAspIleSerGln 656
Qy 1615 GTGAGCTGGAAGCCCTTCGAGGGGAGTGGAGTGTACTCAGATGAGTGCCTCTTC 1674
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 657 ValThrGlnAlaSerIleuArgSerHisIleGlyValValProGlnAspThrValLeuPhe 676
Qy 1675 CATATATCTATTTATTAACCTCTTATATGAAACATCAAGCTTCACTGAGAAAGTG 1734
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 677 AsnAspThrIleAlaGlnAsnIleArgTyrGlyArgValThrAlaGlyAsnAspGlyVal 696
Qy 1735 TATGACATGGCAAAATAGTGGACCTTCATGATGCAATTCCTGCAATGCCCATGATAT 1794
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 697 GluAlaIaIaIaGlnAlaIaGlyIleHisAspAlaIleMetAlaPheProGluGlyTyr 716
Qy 1795 GACACCAAGTAGGGGAGGAGACGACTCAAGCTTCAGAGAGAAAGCAAGAGTAGCA 1854
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 717 ArgThrGlnValGlyGlnArgGlyLeuIleuSerGlyGlyGlnArgProValAla 736
Qy 1855 ATTGCAAGAGCATTGTAAGAGACCCCGCAGTCATCTATGATGAGACTTCTCATCG 1914
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 737 IleAlaArgThrIleLeuValProGlyIleIleLeuLeuAspGlyAlaThrSerAla 756
Qy 1915 TTATATTCGATTACTGAAAGACATATTCTGTGTCATGAAAGATGTGTCAAACACAGA 1974
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 757 LeuAspThrSerAsnGluArgAlaIleGlnAlaSerLeuAlaValCysAlaAsnArg 776
Qy 1975 ACTCTATTTCTTCGACACAGATGTCACAGTGTGATGAGTGAATGATGTC 2034
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 777 ThrThrIleValValAlaHisArgLeuSerThrValValAlaAspGlnIleLeuVal 796
Qy 2035 TTGGATCAGGGTAGGTAAGCGCAAGCTGTGATCCACCATGTTGCTTACCCCTCAT 2094
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 797 IleLysAspGlyCysIleValGluArgGlyArgHisGluAlaLeuLeuSerArg---Gly 815
Qy 2095 AGTATCTATTCAGAAATGTGGCATATACAGAGACCGCTGTGCGAAGCATGATACCCC 2154
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 816 GlyValIyrAlaAspMetTrpGlnLeuGlnGln-----GlnGlyGlnGluThr 832
Qy 2155 AAATGGAGCAAGAAAGAAATATATATCCAA 2187
Db 833 SerGluAspThrIyrProGlnThrMetGluArg 843

RESULT 10
ID AAU00011 standard; Protein; 843 AA.
AC AAU00011;
DT 08-MAY-2001 (first entry)
DE Human ABCB12 transporter.
XX
XX Human; ABCB12 transporter; ATP binding cassette; Alzheimer's disease;
XX human; ABCB12 transporter; ATP binding cassette; Alzheimer's disease;
XX macular degeneration; cystic fibrosis; attherosclerosis;
XX Huntington's disease; Parkinson's disease; mood disorder; panic;
XX antibody; immunogen.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 27..50 /label= Transmembrane_helix_1
XX MISC-difference 254 /note= "Wild-type Trp replaced with Xaa. Encoded by
XX Trp. This substitution is probably due to a misprint in
XX the specification"

```

```

FT Domain 262..285
FT /label= Transmembrane_helix_2
FT Domain 304..326
FT /label= Transmembrane_helix_3
FT Domain 383..405
FT /label= Transmembrane_helix_4
FT Domain 503..521
FT /label= Transmembrane_helix_5
FT Domain 530..548
FT /label= Transmembrane_helix_6
FT Region 623..630
FT /label= Walker_A_motif
FT Region 727..738
FT /label= ATP-binding_cassette_signature
FT Region 739..754
FT /label= Walker_B_motif

XX
XX WO200114547-A1.
XX
XX 01-MAR-2001.
XX
XX 18-AUG-2000; 2000WO-CA00964.
XX
XX 20-AUG-1999; 99US-0161724.
XX
XX 30-AUG-1999; 99US-0151473.
XX
XX 17-AUG-2000; 2000US-0641353.
XX
XX (ACT1-) ACTIVEPASS PHARM INC.
XX
XX Le Bihan S, Wilson C;
XX
XX WPI; 2001-218448/22.
XX
XX N-PSDB; AAS00010.
XX
XX Novel isolated ABCB12 transporter nucleic acid molecule useful as
XX target for developing modulating agents of multidrug resistance and in
XX diagnosis and therapy for Alzheimer's disease and, mood and panic
XX disorders -
XX
XX Claim 12; Fig 2; 97pp; English.
XX
XX The sequence represents a human ATP binding cassette (ABC) transporter
XX family member, ABCB12. The nucleic acid encoding ABCB12 is useful as a
XX target for developing modulating agents of multidrug resistance and as
XX diagnostic and therapeutic tools to control cellular transport in cells
XX of the brain and transport across the blood-brain-barrier and for
XX developing strategies to deliver drugs to the brain, for treating
XX Alzheimer's disease, hyperinsulinaemic hypoglycaemia, cystic fibrosis,
XX attherosclerosis, macular degeneration, adrenoleukodystrophy, prion
XX diseases, Huntington's disease, Parkinson's disease and for treating
XX mood and panic disorders. The nucleic acid is useful for expressing
XX ABCB12, to detect ABCB12 transporter mRNA or genetic alteration in an
XX ABCB12 transporter gene, and to modulate ABCB12 transporter activity and
XX is useful in tissue-typing to identify individuals in forensic
XX biology. The protein is useful as an immunogen to raise anti-ABCB12
XX transporter antibodies and is used to screen for naturally occurring
XX ABCB12 transporter substrates and to screen for drugs or compounds which
XX modulate ABCB12 transport activity. The nucleic acid and protein are used
XX to treat a subject having a disorder characterised by aberrant or
XX unwanted ABCB12 transporter protein or nucleic acid expression or
XX activity. The nucleic acid, protein and antibodies against ABCB12 are
XX useful in screening assays, predictive medicine and in prophylactic and
XX therapeutic treatment. The antibodies are useful to detect and isolate
XX ABCB12, regulate the bioavailability of ABCB12 and to modulate ABCB12
XX transporter activity. A vector expressing the nucleic acid is useful for
XX producing non-human transgenic animals.
XX
XX
XX Sequence 843 AA:
XX
XX
XX Alignment Scores:
XX Pred. No.: 3,08e-105 Length: 843
XX Score: 1218.00 Matches: 283
XX Percent Similarity: 53.44% Conservative: 129

```

Best Local Similarity: 36.71% Mismatches: 251
Query Match: 28.63% Indels: 108
DB: 22 Gaps: 16
AF133659 (1-2345) x AAU00011 (1-843)

QY 10 CTCGGGATGCATTTCTGGCGCTGGCGCGCGCGCTGCTTTGGAAAAAGCCGGCAC 69
DB 136 LeuAlaMetGly1LeuTrp-----1LeuSphaArgHis 146
QY 70 TCC-----GGGATTCGATCCGGCCTTACTCTCTT 102
DB 147 SerProGlyLeuLeuLeuLeuTrpValAlaPheAlaGluAsnLeuAlaLeuVal 166
QY 103 AGCGGCTCAGTCCCGCAGTGGAGGCCACATCACTCGCGCCTGGGAAACCGCTCGAGCC 162
DB 167 SerTrpAsnSerProGlnTrp----- 173
QY 163 TACCAAGATTCACAGATCATTTAAAAAGTATCAGATGGACAGATTGGGAAAGCAATTCA 222
DB 174 -----TrrTrpAlaArgAlaAspLeuGlyGlnGln 183
QY 223 GGACAGCTTCTTAGATGCTGCAAAAGGCTCTCCAGGTATGGCCACTGATAGAAAAAGAGACA 282
DB 184 ValGlnPhe-----SerLeuTrpValLeu----- 191
QY 283 TGTGGCATGTGCATGACGAGAGAGAGACTCCACACA----- 318
DB 192 -----ArgTyrValValSerGlyLeuPheValLeuGlyLeuTrpAlaProGlyLeu 209
QY 319 -----GACCCAAAAAGAGGGTAAAGAGTGTGATATCTGGAAA 357
DB 210 ArgProGlnSerTyrThrLeuGlnValHisGluGluAspGluAspValGluArgSerGln 229
QY 358 ATC-----ATAAAGCAATG 372
DB 230 ValArgSerAlaAlaGlnGlnSerThrTrpArgAspPheGlyArgLeuArgLeuLeu 249
QY 373 CTTTCTTATGTGTGGCCCAAGACAGCCAGATCTTACAGATGAGTTCATTCGCTG 432
DB 250 SerGlyTyrLeu***ProArgGlySerProAlaLeuGlnLeuValLeuLeuLeuLeuLeu 269
QY 433 GGATTTTGGGTGGTCAAAAGCCATGAATATTTGGTTCCTTCATGTTAAATATGCT 492
DB 270 GlyLeuMetGlyLeuGlnLysArgAlaLeuAsnValLeuValProIlePheTyrArgAspIle 289
QY 493 GTCAGACGCTCAACCAAGTGTGGGAAACATGCTGAACCTGAGATGACACA---AAT 549
DB 290 ValAsnLeuLeuThrGlu-----LysAlaProTrpAsn 300
QY 550 ACAGTTGCA---ACCATGGCAACAGAGCTTGTGATGGCTAT-----GGTGTATCAGA 600
DB 301 SerLeuAlaIleTrpThrValThrThrTyrValPheLeuLysPheLeuGlnGlyGlyThr 320
QY 601 GCTGAGAGCTCTTTTAAACGAAGTTCGAAATGCAATTTTGGCAGAGTACCCAGAAAT 660
DB 321 GlySerThrGlyPheValSerAsnLeuArgThrPheLeuTrpIleArgValGlnGlnPhe 340
QY 661 TCATATCCGAAGAATACCCAAAATGCTTTCATCTTCACAACTGATCTGGGTTT 720
DB 341 ThrSerArgValGluLeuLeuLeuPheSerHisLeuHisGluLeuSerLeuArgTrp 360
QY 721 CACCTGAGACAGACAGCGGAGCTTATCTAAGCTATGAGCTATGACAGAGAAACAGGGATAC 780
DB 361 HisLeuGlyArgArgThrGlyGluValLeuArgIleAlaAspArgGlyThrSerSerVal 380
QY 781 AGTTTTCCTGAGTGCCTTTGATTTAACTCTTCCCATGATGTTGAAGTATGCTT 840
DB 381 ThrGlyLeuLeuSerTyrLeuValPheAsnValIleProThrLeuAlaAspIleIle 400
QY 841 GTCAGTGTGTTTGTATTAACAATGC-----GGTGGCCAGTTTGTCTTGTAAACCTT 894
DB 401 -----GlyIleIleTyrPheSerMetPhePheAsnAlaIleTrpPheGlyLeuIleValPhe 418

QY 895 GGAACACTTGGTATACATACAGACATTCACAGTTCACACGCTGGAGAACTAGATT 954
DB 419 LeuCysMetSerLeuTyrLeuThrLeuThrIleValValThrGluTrpArgThrLysPhe 438
QY 955 AGAATGAATAAGAACAAAGCAGATTAATGATGCAGTATAGCTGTATATGACTCTGTG 1014
DB 439 ArgArgAlaMetAsnThrGlnGluAsnAlaThrArgAlaArgAlaValAspSerLeuLeu 458
QY 1015 AATTATGAACCTGAGATATTTAATATGAAGATATGACAGACAGAGATATGATGA 1074
DB 459 AsnPheGluTrpValLysTyrTyrAsnAlaGluSerTyrGluValGluArgTyrArgGlu 478
QY 1075 TTTTGAAGAGTATGAGACTGCTTCATTTGAAAAGTACCTTACTGCTGCTTGAAC 1134
DB 479 AlaIleIleLysTyrGlnGlyLeuGluTrpLysSerSerAlaSerLeuValLeuAsn 498
QY 1135 TTTGTCAAAGTGTCTATTTTACGTGTGCTTTAACAGTATTAATGTCGTCCAGTCAG 1194
DB 499 GlnThrGlnAsnLeuValIleGlyLeuGlyLeuLeuAlaGlySerLeuLeuCysAlaTyr 518
QY 1195 GGAATTGTGGCAGGATACCTTACTGTTGAGATCTAGTAATAGTGAATGACCTCTTTT 1254
DB 519 PheValThrGluGlnLysLeuGlnValGlyAspTyrValLeuPheGlyThrTyrIleIle 538
QY 1255 CAGCTTCATTACCCCTGAACTTCTGTGGAACCTGATATAGACAGACTAGACAAGACTC 1314
DB 539 GlnLeuTyrMetProLeuAsnTrpPheGlyThrTyrTyrArgMetIleGlnThrAsnPhe 558
QY 1315 ATAGATATGAACACCTTGTACTTACTCTACTCAAGTACACCCCAATTAAGACAAATG 1374
DB 559 IleAspMetGluAsnMetPheAspLeuLeuLysGluGluThrGluValLysAspLeuPro 578
QY 1375 ATGGCATCTCCCTTCAGATCACACACAGACACTACCGGCGCTTGTATTAATGTGAT 1434
DB 579 GlyAlaGlyProLeuArgPhe-----GlnLysGlyArgIleGluPheGluAsnValHis 596
QY 1435 TTTGATATCATTTAGGGGCGAAGAAAGTCCTTAGTGAATATCTTTGAAGTCCTCGACA 1494
DB 597 PheSerTyrAlaAspGlyArgGluThrLeuGlnAspValSerPheThrValMetProGly 616
QY 1495 AAGAAAGTGGCATTTGTAGAGGTAGTGGTCAAGGAAAGACAAATATGAGGCTATTA 1554
DB 617 GlnThrLeuAlaLeuValGlyProSerGlyValGlyLysSerThrIleLeuArgLeuLeu 636
QY 1555 TTTGCTCTTATGAGCCCTCAAAAGGATGACATTATCTGTGGTCAAAATATGCAAGAT 1614
DB 637 PheArgPheTyrAspIleSerSerGlyCysIleArgIleAspGlyGluAspIleSerGln 656
QY 1615 GTGAGCCTGGAAAGCCTTCGAGGAGGAGTGGAGTGTACTCAGATGCTGTCTTTC 1674
DB 657 ValThrGlnAlaSerLeuArgSerHisIleGlyValValProGlnAspThrValLeuPhe 676
QY 1675 CATATATCTATTTATTTACACCTCTTATATGAAAACATCACTGCTTACCTGAGAAATG 1734
DB 677 AsnAspThrIleAlaGluAsnIleArgTyrArgValThrAlaGlyAsnAspGluVal 696
QY 1735 TATGAGTGGGAAATTAAGCTGAGCTTCATGATGCAATTTCTGGAATGCCCATGATAT 1794
DB 697 GluAlaAlaAlaGlnAlaAlaGlyIleHisAspAlaIleMetAlaPheProGluGlyTyr 716
QY 1795 GACACCCCAAGTAGGAGAACGAGACTCAAGCTTTCAAGAGAGAGAAAGCAAGATGACA 1854
DB 717 ArgThrGlnValGlyGluArgGlyLeuLysLeuSerGlyGlyGluLysGlnProValAla 736
QY 1855 ATTGCAAGAGCACTTTTGAAGACCCCGCCAGTCACTCTATGATGAAGTACTTTCATCG 1914
DB 737 IleAlaArgThrIleLeuLysValProGlyIleLeuLeuAspGluAlaThrSerAla 756
QY 1915 TTAGATTCGATTACTGAAGACATATTCCTGTGTCATGAGATGAGTGTGCAACACAGA 1974
DB 757 LeuAspThrSerAsnGluArgAlaIleGlnAlaSerLeuAlaLysValCysAlaAsnArg 776

QY 1975 ACTTCTATTTCATGCAACAGATGTCACAGTGGTTGATGCAGATGAATCATTTGTC 2034
 Db 777 ThrThrIleValIValAlaHisArgLeuSerThrValAlaSerAlaAspGlnIleuVal 796
 QY 2035 TTGGATGACGAGTAAGTCCGACCGGTATCCACCATGTTGGTTCGTAACCCCTCAT 2094
 Db 797 IleIysAspGlyCysIleValIgluArgGlyArgHisGluAlaLeuSerArg--Gly 815
 QY 2095 AGTATCTATTGCAAGAAATGTGCATACACAGACGCGGTGGCAAAACCATGATTAACCCC 2154
 Db 816 GlyValIThrAlaAspMetTrpGlnLeuGlnI--GlnGlyGlnGlnGluThr 832
 QY 2155 AATGGGAGCAAGAAAGAAATATATATCCAA 2187
 Db 833 SerGluAspThrLysProGlnThrMetGluArg 843

RESULT 11
 ID AAY08961 standard; Protein; 475 AA.
 XX AAY08961;
 AC AAY08961;
 DT 19-AUG-1999 (first entry)
 XX A. goesypii ORF 1 encoded protein.
 DE
 XX Purine biosynthesis; phosphoribosylpyrophosphate synthetase; KPR1; KPR2;
 KM glutamine-phosphoribosylpyrophosphate amidotransferase; ADE4; GUA1;
 KM IMP-dehydrogenase; GUA1; GMP-synthetase; GUA2; PRRP synthetase; mouth;
 KM riboflavin; vitamin B2; treatment; inflammation; throat; mucosal layer;
 KM skin disorder.
 XX Aeshbya goesypii.
 OS
 PN EP927761-A2.
 XX
 PD 07-JUL-1999.
 XX
 PF 08-DEC-1998; 98EP-0123331.
 XX
 PR 23-DEC-1997; 97DE-1057755.
 XX
 PA (BAD1) BASF AG.
 XX
 PI Hoeffken HW, Jimenez A, Pompejus M, Revuelta Doval JL;
 PI Santos Garcia MA, Seubberger H;
 XX
 DR WPI; 1999-373465/32.
 DR N-PSDB; AAX78080.
 XX
 XX Aeshbya goesypii purine biosynthesis enzymes and related genes for
 PT production of riboflavin
 XX
 PS Example 3; Page 22-24; 48pp; German.
 XX
 CC This invention describes novel enzymes involved in purine biosynthesis
 CC in Aeshbya goesypii. The enzymes described in the invention include
 CC Phosphoribosylpyrophosphate (PPRP) synthetases KPR1 and KPR2
 CC glutamine-phosphoribosylpyrophosphate (PPRP) amidotransferase (ADE4)
 CC IMP-dehydrogenase (GUA1) and a GMP-synthetase (GUA2). The nucleic acid
 CC sequences encoding the enzymes involved in purine biosynthesis are used
 CC to genetically modify microorganisms, for the production of riboflavin.
 CC Riboflavin, also known as vitamin B2, is essential in humans and animals
 CC and is used to treat inflammation of the mouth and throat mucosal layer
 CC as well as inflammation in skin disorders. This sequence represents the
 CC protein product of an open reading frame which flanks the ADE4 gene which
 CC encodes a glutamine PRP amidotransferase.
 XX
 SQ Sequence 475 AA;
 Alignment Scores: 8.44e-104 Length: 475
 Pred. No.: 1201.50 Matches: 234
 Score:

Percent Similarity: 71.74%
 Best Local Similarity: 50.87%
 Query Match: 28.24%
 DB: 20 Gaps: 2

AF133659 (1-2345) x AAY08961 (1-475)

QY 757 ATTGACAGAGAAACAAGGGGTATCAGTTTGTCTGAGCTTTGGTATTATTCCTT 816
 Db 1 MetAspArgGlyCysLysGlyIleSerTyrValLeuSerIleMetValPheHisIleIle 20
 QY 817 CCATCATGTTTGAAGTATGATCTTGTGAGTGTGTTTGTATTACAATGCGGTGCCAG 876
 Db 21 ProIleThrPheGlnIleSerMetValCysGlyIleLeuThrTyrGlnPheGlyAlaSer 40
 QY 877 TTTGCTTTGTAACCCCTTGAGACCTTGATACATACACACATTCACAGTTCAGTCA 936
 Db 41 PheAlaIleIleThrPheSerThrMetLeuLeuTyrSerIlePheThrPheArgThrThr 60
 QY 937 CCGTGGAGAACTAGATTATTAAGTAATGACAAAGCAGATATGATGACGATATGCT 996
 Db 61 AlaTrpArgThrArgPheArgAspAlaAsnLysAlaAspAsnLysAlaAlaSerVal 80
 QY 997 GCTATGACTCTGCTGTAATTATGAAACTGTGAAGTATTATTAATGAAGATATGAA 1056
 Db 81 AlaLeuAspSerLeuIleAsnPheGluAlaValLysTyrPheAsnGlnLysTyrLeu 100
 QY 1057 GCACAGAGATATGATGATTTTGAAGACCTATGAGACTGCTCATTTGAAGAACTCT 1116
 Db 101 AlaAspLysTyrIleThrSerLeuMetLysTyrArgAspSerGlnIleLysValSerGln 120
 QY 1117 ACTCTGGCTATGCTGAACCTTGTGCAAGTGCTATTGTCGGTTTAACAGCTATA 1176
 Db 121 SerLeuAlaPheLeuAsnThrGlyGlnAsnLeuIlePheThrAlaLeuThrAlaMet 140
 QY 1177 ATGGTCTGCCAGTCAAGGAATTGTGGCAGTACCTTACTGTGGAGATCTTAATG 1236
 Db 141 MetTyrMetAlaCysAsnGlyValMetGlnGlySerLeuThrValGlyAspLeuValLeu 160
 QY 1237 GTGAATGCACTGCTTTTTCAGCTTTTCACTTACCCCTTAACCTTGTGGAACTGATATGA 1296
 Db 161 IleAsnGlnLeuValPheGlnLeuSerValProLeuAsnPheLeuGlnSerValTyrArg 180
 QY 1297 GAGACTAGACAAGCACTCATAGATATGAACACTGTTTACTTACTCAAGGTAGACACC 1356
 Db 181 AspLeuLysGlnSerLeuIleAspMetGlnSerLeuPheLysLeuGlnLysAsnGlnVal 200
 QY 1357 CAATTAAGACAAGATGATGCGATCTCCCTTCAGATCAACACAGACAGTACCGTG 1416
 Db 201 ThrIleLysAsnSerProAsnAlaGlnAsnLeuProIle--HisLysProLeuAspIle 219
 QY 1417 GCCTTTGATTAATGTGATTTTGAATATACATTGAGGCGCAAAATCCTTATGGAATATCC 1476
 Db 220 ArgPheGlnAsnValThrPheGlyTyrAspProGluArgGlyIleLeuAsnValSer 239
 QY 1477 TTGGAAGTCCCTGACAGAAAGAAAGTGGCATGTGAGAGTGTGGTTCAGGGAAGC 1536
 Db 240 PheThrIleProIleArgLysMetLysThrAlaIleValGlyProSerGlySerGlyLysSer 259
 QY 1537 ACAATGATGAGCTATTAATTTGCTTCTTATGAGCCTCAAAAGGTAGACATTAATCTTGTCT 1596
 Db 260 ThrIleLeuLysLeuValPheArgPheTyrGluProGlnGlnGlyArgIleLeuValGly 279
 QY 1597 GGTCAAAATATTAACAATGTGAGCCTTGAAAGCCTTCGAGAGGCACTGGAAGTGTACT 1656
 Db 280 GlyThrAspLysPheAspLeuSerLeuArgLysAlaIleGlyValPro 299
 QY 1657 CAGATGCTGCTCTTCAATATAATCTATTATTAACAACCTCTTATGGAAGCAATCATG 1716
 Db 300 GlnAspThrProLeuPheAsnAspThrIleTrpGluAsnValLysPheGlyAsnIleSer 319
 QY 1717 GCTTCACTGAGAGAGATGATGCAAGTGGCAAAATTAAGTCTGACTGATGATGCAATTTCT 1776

Db	320	Server	serp	serp	gull	leu	leu	ala	leu	gln	leu	thr	tyr	ser	leu	leu	gln	333
Qy	1777	CGAATGCCACATGGATATGACACCCAAAGTGGGGAACGAGAACTCAAGCTTTACAGAGAA	18368															
		:::																
Db	340	AsnleuProlysglYlaSerthrValaIglYlaIrglYleuMetIleSerIglY	359															
Qy	1837	GAAGAAGCAAGAGTNGCAATTTGCAAGAGCCATTTTGAAGAGCCGCCACTCACTACTCAT	18966															
Db	360	GluYlsglnIargYleuIalaIlealaIglYValIleuYlserIalaIroYleuMetPhe	379															
Qy	1897	GATGAGACCTACTTCATGCTTGAAGTTCATGATTCACGAAGACACTATTTCTGGTCCATGAAG	19566															
Db	380	AspIgluIalthrSerIalaIeudSerThrIlethIglYlalaIleuIenIsthrIleGln	399															
Qy	1957	GAT-----GTGTCAAACACAGAACTTCTATTTTCATTCACACAGATTCGAACAGT	2010															
Db	400	GlnAsnPheSerSerAsnSerIylsThrSerValIlyrValaIalIstIargYleuIargThrlle	419															
Qy	2011	GTTCATGACAGATGAATCAATCTGCTTGATGACAGAGGTAGTGGACGAGTGTCCAC	20707															
Db	420	AlaAspIalaAspIylsIleIleValIleGlnIglYlSerValaIrgIluIuIylThrIle	439															
Qy	2071	CATGTTTCTGCTTGTCAACCTCATAGTATCTATTCAGAAATGGCATCAACAGACAGC	2130															
Db	440	SerSerIleuIeulaIaSerGlnIglYSerIeutyIArgIglYleuThrAspIleIglYlAsn	459															
RESULT 12																		
AA5454	AA5454	standard; Protein; 574 AA.																
XX	AA54454;																	
XX	25-APR-2000	(first entry)																
DE	Amino acid sequence of a human transport protein homologue HTPH-1.																	
XX																		
KM	Human; transport protein homologue; HTPH-1; ABC transporter; cancer;																	
KM	reproductive disorder; adenocarcinoma; leukemia; lymphoma; melanoma;																	
KM	myeloma; sarcoma; testicular carcinoma; prolactin production; infertility;																	
KM	ovarian hyperstimulation syndrome; uterine fibroid; Peyronie's disease;																	
KM	gynecomaastia; copper metabolism disorder; Menke's disease;																	
KM	Wilson's disease; Ehlers-Danlos syndrome type IX.																	
OS	Homo sapiens.																	
XX																		
XX	Key	Location/Qualifiers																
FT	Modified-site	26	/note=	"potential phosphorylation site"														
FT	Modified-site	73	/note=	"potential phosphorylation site"														
FT	Modified-site	74	/note=	"potential phosphorylation site"														
FT	Modified-site	89	/note=	"potential phosphorylation site"														
FT	Modified-site	126	/note=	"potential phosphorylation site"														
FT	Modified-site	179	/note=	"potential phosphorylation site"														
FT	Modified-site	194	/note=	"potential glycosylation site"														
FT	Modified-site	230	/note=	"potential phosphorylation site"														
FT	Modified-site	330	/note=	"potential glycosylation site"														
FT	Modified-site	337	/note=	"potential phosphorylation site"														
FT	Modified-site	355..362	/note=	"potential														

FT	Region	/note= "potential phosphorylation site"
FT		459..473
FT	Modified-site	/note= "ABC transporter signature sequence"
FT		460
FT	Modified-site	/note= "potential phosphorylation site"
FT		487
FT	Modified-site	/note= "potential phosphorylation site"
FT		491
FT	Modified-site	/note= "potential phosphorylation site"
FT		507
FT	Modified-site	/note= "potential glycosylation site"
FT		563
FT	Modified-site	/note= "potential phosphorylation site"
PX		
PX	W0200003015-A2.	
PX		
PD	20-JAN-2000.	
PX		
PF	08-JUL-1999;	99MO-US15482.
PX		
PR	10-JUL-1998;	98US-0113427.
PX		
PA	(INCY-) INCYTE PHARM INC.	
PX		
PI	Hillman JL, Yue H, Reddy RM, Gorgone GA, Corley NC, Azimzai Y,	
PI	Patterson C, Baughn MR;	
DR	WPJ; 2000-160925/14.	
DR	N-PSDB; AA245823.	
XX		
PT	New human transport protein homologs, polypeptides and nucleic acids	
PT	useful for diagnosing, treating or preventing cancer, reproductive	
PT	disorders and copper metabolism disorders -	
XX		
PS	Claim 1; Page 68-69; 75pp; English.	
XX		
CC	The present sequence represents a human transport protein homologue,	
CC	designated HTPH-1. HTPH-1 is a member of the ABC transporter family.	
CC	The antagonists of the protein are useful for treating or preventing	
CC	cancer and reproductive disorders such as adenocarcinoma, leukemia,	
CC	lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, etc., and	
CC	disorders of prolactin production, infertility, ovarian, hyperstimulation	
CC	syndrome, uterine fibroids, Peyronie's disease, gynecomastia, etc.. The	
CC	polynucleotide, antibodies, agonists, complementary sequences or	
CC	vectors are also useful for treating or preventing the diseases. The	
CC	polypeptide is also useful for treating or preventing a copper	
CC	metabolism disorders such as Menke's disease, Wilson's disease and	
CC	Ehlers-Danlos syndrome type IX. The antibodies are also useful for	
CC	diagnosing disorders characterized by expression of the human transport	
CC	protein homologue.	
XX		
SQ	Sequence 574 AA;	
Alignment Scores:		
Fred. No.:	5.03e-102	Length: 574
Score:	1183.00	Matches: 250
Percent Similarity:	62.12%	Conservative: 114
Best Local Similarity:	42.66%	Mismatches: 194
Query Match:	27.81%	Indels: 28
DB:	21	Gaps: 9
AF133659 (1-2345) x AAY54454 (1-574)		
OY	CTGGGATTTTGGGTGTGCAGAGGCCATGAATTATGTTCCCTTCATGTTTAATAAT	489
Db	1 MetcyleumetctgValaIargalalaLeuaInvalProlleheyrarig---	19
OY	490 GCTGTAGACAGCCTCAACAGATGTGGGAAACATGCTGAACCTG-----AGTAGTCA	543
Db	20 -----AsnlleValasleuHeuthrGlulysala	29
OY	544 CCA---AATACAGTTGCACCATGCGAACAGCA-----GTTTGATTTGGC	585

Dh 30 ProTirpaNserLeuAlaTirThValThSerTyValPheLeuLysPheLeuGlnGly 49
Qy 586 TATGTGTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTGCAAAATGCAGATATTTGGC 645
Dh 50 GlyGlyThrGlySerThrGly-----PheValSerAenLeuAlaGThrPheLeuTriPle 67
Qy 646 AAGGTAGCCCAAGATTCAATCCGAAGATACCCAAAATGCTTTCTTCATCTTCACAAAC 705
Dh 68 ArgValGlnGlnPheThrSerArgValGluLeuLeuIlePheSerHisLeuHisGlu 87
Qy 706 CTGATCTGGGTTTTCACCTGAGCAGACAGGGAGCTTATCATAGGCTATTGACAGA 765
Dh 88 LeuSerLeuAlaGThrPheLeuGlyArgArgThrGlyGluValLeuArgGlyLeuAlaAspArg 107
Qy 766 GAACAAGAGGGTATCAGTTTGTCTGAGTCTTGATTTAAATCTTCCATCATG 825
Dh 108 GlyThrSerSerValThrGlyLeuLeuSerTyLeuValPheAenValIleProThrLeu 127
Qy 826 TTGAAAGTATGCTTTCAGTGTGTTTGTATTACAAATCC-----GGTCCCAAGTT 879
Dh 128 AlaAspIleIleIle-----GlyIleIleTyPheSerMetPhePheAsnAlaTirPhe 145
Qy 880 GCTTTGAACCTTGGACAACCTGTGATACACAGATTCAGATTCGATGCACACGG 939
Dh 146 GlyLeuIleValPheLeuCysMetSerLeuTyLeuThrIleuThrIleValValThrGlu 165
Qy 940 TGGAGAACTAGATTAGAAATGAAATGAAACAAACAGATATGACAGTAATGTGCT 999
Dh 166 TTrpArgThrLysPheArgArgAlaMetCAsnThrGlnGlnAenAlaThrArgAlaArgAla 185
Qy 1000 ATGACTCATCTGCTGAATTAAGAACTGTGAAAGTATTTAATAAGAAATATGAAACA 1059
Dh 186 ValAspSerLeuLeuAenPheGluThrValLysTyTrpAenAlaGluSerTyGluVal 205
Qy 1060 CAGAGATATGATGATTTTGAACAGCATAGACGTCTTATGAAAAGTACCTTACT 1119
Dh 206 GluArgTyArgGluAlaIleIleLysTyGlnGlyLeuGluLurPylsSerSerAlaSer 225
Qy 1120 CTGGCTATGCTGAACCTTGTGCAAGTGTCTATTTCAGTGTGCTTAAACAGTATATG 1179
Dh 226 LeuValLeuLeuAenGlnThrGlnAenLeuValIleGlyLeuGlyLeuLeuAlaGlySer 245
Qy 1180 GTGCTGCCACGTACAGGAATTTGGACAGTACCTTACTGTGGAGATCTAGTAATGTG 1239
Dh 246 LeuLeuCysAlaTyPheValThrGluGlnLysLeuGlnValGlyAspTyValLeuPhe 265
Qy 1240 AATGACTGCTTTTTCAGCTTTTCAATACCTGTAACCTTCTCGGAATCTATATAGAG 1299
Dh 266 GlyThrTyTrpIleIleGlnLeuTyMetProLeuAenThrPheGlyThrTyTrpArgMet 285
Qy 1300 ACTGACACAGCACTCATGATATGAAACCTTGTTCCTACTACAGATAGACACCCAA 1359
Dh 286 IleGlnThrAenPheIleAspMetGluAenMetPheAspLeuLysGluGluThrGlu 305
Qy 1360 ATTAAGACAAAGTATGCGATCTCCCTTCAGATCAACACACAGACGTAACCGTGCC 1419
Dh 306 ValLysAspLeuProGlyAlaGlyProLeuArgPhe-----GlnLysGlyArgGlu 323
Qy 1420 TTTCATATATGTCATTTTGAATACATTGAGGGCCAAAGTCTTACTGCAATATCTCTT 1479
Dh 324 PheGluAenValHisPheSerTyAlaAspGlyArgGluThrGlnAenValSerPhe 343
Qy 1480 GAAGTCCCTGAGGAAGAAAGTGGCCATTTGAGAGTGTGGTACAGGGAAAGACA 1539
Dh 344 ThrValMetProGlyGlnThrLeuAlaLeuValGlyProSerGlyAlaGlyLysSerThr 363
Qy 1540 ATAGTGAAGCTATTATTTGCTTCTTATGAGCCTCAAAAGGTATGACTTATCTGTGCT 1599
Dh 364 IleLeuArgLeuLeuPheArgPheTyAspIleSerSerGlyCysIleArgGlyLeuAspGly 383
Qy 1600 CAAATATACAAAGTGTGAGCTTGAAAGCTTTGAGGGCAGTGGAGAGTGTACTCAG 1659
Dh 384 GlnAspIleSerGlnValThrGlnAlaSerLeuArgSerHisIleGlyValValProGln 403

Qy 1660 GATGCTGTCTCTTCATTAATACCTTATTAACACCTTATATGMAACATCAGTCT 1719
Dh 404 AspThrValLeuPheAenAspThrIleAlaAspAsnIleArgTyArgValThrAla 423
Qy 1720 TCACCTGAGGAAGTGTATGACAGTGGCAAAATTTGCTGAGCTTATGATATGCAATCTTCCA 1779
Dh 424 GlyAsnAspGluValGluAlaAlaGlnAlaAlaGlyIleHisAspAlaIleMetAla 443
Qy 1780 ATGCACATGATATGACACCAAGTAGGGGAGCAGAGACTCAAGCTTCAGAGAGAA 1839
Dh 444 PheProGluGlyTyArgThrGlnValGlyLurArgGlyLeuLysLeuSerGlyGlyGlu 463
Qy 1840 AAGCAAGATAGCAATTTGCAAGAGCCATTTTGAAGAGCCCCCATCATCTATGAT 1899
Dh 464 LysGlnArgValAlaIleAlaIleArgThrIleLeuLysAlaProGlyIleIleLeuLeuAsp 483
Qy 1900 GAAGCTTACTTATCGTATGATTTGATTAAGAGACTATTTCTGGTCCATGAAGAT 1959
Dh 484 GluAlaThrSerAlaLeuAspThrSerAsnGluArgAlaIleGlnAlaSerLeuAlaLys 503
Qy 1960 GTGGTAAACAGAACTTCTATTTTCATTGACACAGATTCACAGGTGTGATGCA 2019
Dh 504 ValCysAlaAsnArgThrThrIleValAlaAlaHisArgLeuSerThrValValAsnAla 523
Qy 2020 GATGAATCATTTGCTTGGATCAGGTAAAGTACCGAAGCTGTACCCCATGTGTTTG 2079
Dh 524 AspGlnIleLeuValIleLysAspGlyCysIleValGluArgGlyArgHisGluAlaLeu 543
Qy 2080 CTTCCTTAACCTCATATGATCTTATTCAGAAATGTGGCATACACAGACCCGTGTGCAG 2139
Dh 544 LeuSerArg---GlyGlyValTyAlaAspMetTrpGlnLeuGlnGlnGlnGlu 562
Qy 2140 AACCATGATTAACCCCAA 2157
Dh 563 ThrSerGluAspThrLys 568
RESULT 13
ID ABB71210 standard; Protein; 866 AA.
XX ABB71210;
AC ABB71210;
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster polypeptide SEQ ID NO 40422.
XX Drosophila, developmental biology; cell signalling; insecticide;
XX Drosophila, developmental biology; cell signalling; insecticide;
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2001; 2000US-191637P.
XX 23-MAR-2001; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-65660/75.
XX DR N-PSDB; ABL15313.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 40422; 21bp + Sequence Listing; English.

Db 813 LeuLeuAsnLeuAlaGluGlnSerGlyGlySerAspAsnGlyAlaSerAlaGlu 832
QY 2176 AATATATCCAAAGAGAGAGAAAGAAAG 2202
Db 833 -----SerglySerGlyLysArgArg 839
RESULT 14
AAM39313
ID AAM39313 standard; Protein; 571 AA.
XX
AC AAM39313;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2458.
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia.
XX
OS Homo sapiens.
XX
PN MO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-048725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
XX
N-PSDB; AAI58469.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2458; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 571 AA;

Alignment Scores:
Pred. No.: 3,51e-101 Length: 571
Score: 1174.00 Matches: 248
Percent Similarity: 62.26% Conservative: 115
Best Local Similarity: 42.54% Mismatches: 192
Query Match: 27,608 Indels: 28
DB: 22 Gaps: 9
AF133659 (1-2345) x AAM39313 (1-571)
QY 439 TTGGGTGGTGCAGAAAGCCATGAAATATTGTGCTTCCTCATGTTAAATAGCTGAGAC 498
Db 1 MetGlyLeuGluAlaLeuAsnValLeuAlaProIlePheTyrArg----- 16
QY 499 AGCCTCAACAGATGTCGGAAACATGCTGACCTG-----AGTAGACACCA-----AAT 549
Db 17 -----AsnIleValAsnLeuLeuThrgLysAsnAlaProTTrpAsn 29
QY 550 ACAGTTCAGAACCATGCGAACAGCA-----GTTCTGATTGGCTATGGTGA 594
Db 30 SerLeuAlaTrpThrValThrSerTyrValPheLeuLysPheLeuGlnGlyGlyThr 49
QY 595 TCAGAGAGCTGGAGCTGCTTTTAAAGAACTTCGAAATGCAAGTATTGGCAAGTACC 654
Db 50 GlySerThrGly-----PheValSerAsnLeuArgThrPheLeuTrpIleArgValGln 67
QY 655 CAGAAATTCAGATCCAGAGAAATAGCCAAATAATGCTTCATTCACAACTGAGATCTG 714
Db 68 GlnPheThrSerArgArgValGluLeuLeuIlePheSerHisLeuHisGluLeuSerLeu 87
QY 715 GGTGTTTCACCTGAGACAGACAGCGAGCTTATATCAAGCTATTGACAGAGAACAGG 774
Db 88 ArgTrpHisLeuGlnYargTrgTrgGlyValLeuArgIleAlaAspArgGlyThrSer 107
QY 775 GGTATAGTTTGTCTGTAGTGTCTTGGATTTAACTTCTCCATCATGTTTGAAGT 834
Db 108 SerValThrGlyLeuLeuSerTyrLeuValPheAsnValIleProThrLeuAlaAspIle 127
QY 835 ATGCTTGTCAGTGGTGTGTTTGTATTACAAATGC-----GGTCCCAAGTTGCTGGTA 888
Db 128 IleIle-----GlyIleIleTyrPheSerMetPhePheAsnAlaTrpPheGlyLeuIle 145
QY 889 ACCCTTGGAACCTTGTTGATACATACAGCATTCACAGTTGACAGTGCACAGGTGAGACT 948
Db 146 ValPheLeuCysMetSerLeuTyrLeuThrLeuThrIleValValThrGluTrpArgThr 165
QY 949 AGATTGATGAATGAATGAACAAAGCAGATTAATGATGACAGTAACTGCTATAGACTCA 1008
Db 166 LysPheArgArgAlaMetAsnThrGlnGluAsnAlaThrArgAlaArgAlaValAspSer 185
QY 1009 CTGCTGAATTAAGAACTGTCGAAGTATTTTAATGAAGAATGACAGACAGATAT 1068
Db 186 LeuLeuAsnPheGluThrValIleTyrTyrAsnAlaGluSerTrgIleValGluArgTyr 205
QY 1069 GATGATTTTGAAGAGTATGAGACTGCTTATGAAAGTAAAGTACTGCTGAGTAAAG 1128
Db 206 ArgGluAlaIleIleLeuTyrGlnGlyLeuGluTrpLysSerLeuValLeuValLeu 225
QY 1129 CTGAACCTTGTGCAAGAGCTATTTTCACTGCTGTTTAAAGCTATTAATGCTGCGCC 1188
Db 226 LeuAsnGlnThrGlnAsnLeuValIleGlyLeuGlyLeuLeuAlaGlySerLeuLeuCys 245
QY 1189 AGTCAGGGAATGTGCGACAGTACCCTTACTGTGAGATCTGATGATGATGATGACTG 1248
Db 246 AlaTyrPheValThrGlnGlnLysLeuGlnValGlyAspTyrValLeuLeuGlyThrTyr 265
QY 1249 CTTTTCACCTTCAATTAACCTGCAACTTTCGGAACCTGATATGAGAGACTAGACAA 1308
Db 266 IleIleGlnLeuTyrMetProLeuAsnTrpPheGlyThrTyrTyrArgIleGlnThr 285
QY 1309 GCACCTATGATGAAACACCTTGTATTACTTCACTCAAGTACAGCAACCAATTAAGAC 1368

```

Db      286 Asnphel1easpmecGluasnmecPheaspleuLeuylsGluGluThrGluValIysasp 305
QY      1369 AAGAGTAGGAGCTCCCTTCAGATTCACACGACAGACGCTACCGCTTGATAT 1428
Db      306 LeuproglyalaglyProleuarkphe-----GlnlysglyAxyIleGluPheGluasn 323
QY      1429 GTGCAATTTGATATACATTGAGGCGCAGAAAGTCTTAAGTGAATATCCCTTGAACTCCCT 1488
Db      324 ValHisPheSerTyrAlaaspGlyargGluThrLeuGlnaspValSerPheThrAlaMet 343
QY      1489 GCAGAGAAAGAAAGTGGCCATTGTAGAGGTATGGGTCAAGGAAACACATATAGTAGG 1548
Db      344 ProglyGlnThrLeuAlaLeuValIglyProSerGlyAlaGlyLysSerThrIleLeuark 363
QY      1549 CTATATATTCGCTTATAGAGCTCAAAAGGTAGCATATTTATCTTGCTGGTCAAAATATA 1608
Db      364 LeuLeuPheArgPheTyrAspIleSerSerGlyCysIleArgIleAspIlyGlnaspIle 383
QY      1609 CAAGATGTGAGCCTTGAAAGCCTTGCGAGGCGAGCGAGTGGAGTGACTCAGAGATCTGTC 1668
Db      384 SerGlnValThrGlnAlaSerLeuarkSerHisIleGlyValValProGlnaspThrVal 403
QY      1669 CTCCTTCATATATCTATTTATTTATCAACTCTTATATGAAACATGATGCTTCACCTGAG 1728
Db      404 LeupheAsnAspThrIleAlaAspAsnIleArgTyrGlyArgValThrAlaGlyAsnAsp 423
QY      1729 GAAGTGTATGAGTGGCAAAATTAAGTGAAGTCTTCATGATGCAATCTTCGAATGCAACT 1788
Db      424 GluValGluAlaAlaAlaGlnAlaAlaGlyIleHisAspAlaIleMetAlaPheProGlu 443
QY      1789 GGATATGACACCCCAAGTAGAGGAGACGAGACTCAAGCTTTCAGAGAGAAAGCAAGA 1848
Db      444 GlyTyrArgThrGlnValIglyIleuArgIleuylsleuSerGlyGlyGluylsGlnArg 463
QY      1849 GTAGCAATTTGCAAGACCATTTTGAAGACCCCAAGTCATATCTTATGATGAGCTACT 1908
Db      464 ValAlaIleAlaIleArgThrIleuLeuylsAlaProGlyIleIleLeuLeuAspGluAlaThr 483
QY      1909 TCATGTTAGTATGATTCATTAAGTGAAGACTATTTCTTGCGCATAGAGATGTGTCAAA 1968
Db      484 SerAlaLeuAspThrSerAsnGlnIleArgAlaIleGlnAlaSerLeuAlaIleValCysAla 503
QY      1969 CACAGAACTCTATTTTTCACACAGATTTGTCAACAGTGGTGGTGCAGATGAAATC 2028
Db      504 AsnArgThrThrIleValValAlaHisArgLeuSerThrIleValAlaAsnAlaAspGlnIle 523
QY      2029 ATTGCTTGATGATGAGGTAGGATGAGCCGCAACCTGTGATCCCACTGATGTTGCTTGCTAAC 2088
Db      524 LeuValIleIleLysAspGlyCysIleValIglyuArgGlyArgHisGluAlaLeuLeuSerArg 543
QY      2089 CCTCATAGTATCTATTCGAATGTGGCATACACAGAGACCGGTGCGAGAACATGAT 2148
Db      544 ---GlyGlyValTyrAlaAspMetTyrGlnLeuGlnGlnGlyGlnGluThrSerGlu 562
QY      2149 AACCCCAAA 2157
Db      563 AspThrLys 565

```

RESULT 15

AAAM41099
ID AAAM41099 standard; Protein; 591 AA.

XX AAAM41099;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6030.

XX Human; noctropic; immunosuppressant; cyrostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

```

KW      1 leukemia.
XX
OS      Homo sapiens.
XX
PN      W0200153312-A1.
XX
PD      26-JUL-2001.
XX
PF      26-DEC-2000; 2000WO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-0552317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
XX      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI      Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR      WPI: 2001-442253/47.
DR      N-PSDB; AA160255.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
PT      such as central nervous system injuries -
XX
XX      Example 2; SEQ ID NO 6030; 10078pp; English.
XX
CC      The invention relates to human nucleic acids (AA157798-AA161369) and
CC      the encoded polypeptides (AAAM38642-AAAM42213) with noctropic,
CC      immunosuppressant and cyrostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      localised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      activation of the activities such as: immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemia and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX
SO      Sequence 591 AA;

```

Alignment Scores:

Pred. No.: 7.14e-99 Length: 591
Score: 1149.50 Matches: 247
Percent Similarity: 61.14% Conservative: 118
Best Local Similarity: 41.37% Mismatches: 203
Query Match: 27.02% Indels: 29
DB: Gaps: 10

AF133659 (1-2345) x AAAM41099 (1-591)

```

QY      400 CCAAGTCTACAGAGCTAGAGTTGCAATTTGCTGGGATTTTGGTGGTCAAGAGCCATG 459
Db      7 ProAlaLeuGlnLeuGluValLeuIleCysLeuIleuLeuGlyLeuGluArgAlaLeu 26
QY      460 AATATGTGTTCCCTTCATGTTAAATATGCTGTAGACAGCTCAACACAGATCCGGA 519
Db      27 AsnValLeuAlaProIlePheTyrArg----- 35
QY      520 AACATGCTGAACCTG-----AGTAGACACCA--AATACAGTTGCACCATGCAACA 570
Db      36 AsnIleValAsnLeuLeuThrGluAsnAlaProTyrAsnSerIleuAlaIleThrValThr 55

```


QY 571 GCA-----GTTCTGATGGCTATGCTGATCAAGAGCTGCTTTT 615
 Db 56 SerTyrValPheLeuLysPheLeuGlnGlyGlyThrGlySerThnGly-----Phe 73
 QY 616 TTTAACGAGATTCCGAATGCGATTTGGCAAGTACCCGAAATTCATCCGAAGATA 675
 Db 74 ValSerLeuArgThrPheLeuTrpIleArgValGlnGlnPheTrnSerArgVal 93
 QY 676 GCCAAATGCTTTCATCTTCACAACTGATCGGGTTTCACCTGAGAGACAG 735
 Db 94 GluLeuLeuIlePheSerHisLeuHisGluLeuSerLeuArgTrpHisLeuGlyArgArg 113
 QY 736 ACCGAGCTTATCTAAAGCTATTGACAGAGAAACAAGGGATCAAGTTTGCTGAGT 795
 Db 114 ThrGlyValValLeuArgIleAlaAspArgGlyThrSerSerValThnGlyLeuLeuSer 133
 QY 796 GCTTTGATTTATCTTCTCCATCATGTTGAAGTGAAGTGTTCAGAGTGTTTG 855
 Db 134 TyrLeuValPheAsnValIleProThrLeuAlaAspIleIleIle-----GlyIleIle 151
 QY 856 TATTAACAATGTC-----GGTGGCCAGTTTGCTTGGTAACCTTGAAACCTTGATCA 909
 Db 152 TyrPheSerMetPhePheAsnAlaTrpPheGlyLeuIleValPheLeuCysMetSerLeu 171
 QY 910 TACACAGCATTCACAGTGCAGTGCACACGCTGAGAACTAGATTAGAAATGAATGAC 969
 Db 172 TyrLeuThrLeuThrIleValIleValThnGlyTrpArgThrLysPheArgArgAlaMetAsn 191
 QY 970 AAGACAGATTAATGATGCAAGGTAATGCTGCTATGACTGCTGCAATTATGAACGTGTG 1029
 Db 192 ThrGlnLeuAsnAlaThrArgAlaArgAlaValAlaAspSerLeuLeuAsnPheGluThrVal 211
 QY 1030 AAGTATTTTATTAATGAAGATGATGAAGCACAGATPAGATTTTGAACAGTAT 1089
 Db 212 LysTrpTyrAsnAlaGluSerTyrGluValGluTrpArgValAlaIleLeuLysTyr 231
 QY 1090 GAGACTGCTTCAATGAAAGTACTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
 Db 232 GlnGlyLeuGlnTrpLysSerSerAlaSerLeuValLeuLeuAsnGlnThnGlnAsnLeu 251
 QY 1150 ATTTTCAGTGTGCTTTTACAGCTATAATGCTGCTCCGACGTCAGGAAATTTGGCAGGT 1209
 Db 252 ValIleGlyLeuGlyLeuLeuAlaGlySerLeuLeuCysAlaTyrPheValThnGlnGln 271
 QY 1210 ACCCTTACTGTTGAGATCTGATATGCTGATGATGATGATGATGATGATGATGATGAT 1269
 Db 272 LysLeuGlnValGlyAspTyrValLeuPheGlyThrTyrIleIleGlnLeuTyrMetPro 291
 QY 1270 CTGAACCTTCTGGAAGCTGTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1329
 Db 292 LeuAsnTrpPheGlyThrTyrTyrArgMetIleGlnThnAsnPheIleAspMetGluAsn 311
 QY 1330 TTGTTTACTGACTCAAGGATGACACCCAAATTAAGAACAAGTGAATGCTTCCCTT 1389
 Db 312 MetPheAspLeuLeuLysGlyThrGlnValLysAspLeuProGlyAlaGlyProPhe 331
 QY 1390 CAGATCACACACAGACAGACGCTGAGCTTGTATATGTCATTTTGAATATGATGAG 1449
 Db 332 ArgPhe-----GlnLysGlyArgIleGlnPheGlnAsnValHisPheSerTyrAlaAsp 349
 QY 1450 GGCAGAAAGCTCTTAGTGAATATCTTTGAAGTCCCTGACGAGAAAGAAAGTGGCATT 1509
 Db 350 GlyArgGlnTrpLeuGlnAspValSerPheThrValMetProGlyGlnThnLeuAlaLeu 369
 QY 1510 GTAGAGAGTGTGGGTCAAGGAAAGACAAATAGTGAAGGCTATTATTGCTTCTATGAG 1569
 Db 370 ValGlyProSerGlyAlaGlyLysSerThrIleLeuArgLeuLeuPheArgPheTyrAsp 389
 QY 1570 CCTCAAAAGGTACATTTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629
 Db 390 IleSerSerGlyCysIleArgIleAspGlyGlnAspIleSerGlnValThnGlnAlaLeu 409

QY 1630 CTTGCGAGGCA--GTGGAGTGTACTCTCAGAGATGCTGTCTTCTTCAATAACTATT 1686
 Db 410 PheArgPheSerHisTrpGluLeuCysProLysAspThrValLeuPheAsnAspThrIle 429
 QY 1687 TATTACAACTCTTATATGGAACATCAGTCTTCACTGAGAGAGTATGCAAGTGCA 1746
 Db 430 AlaAspAsnIleArgTyrGlyArgValThrAlaGlyAsnAspGluValGluAlaAla 449
 QY 1747 AAATGATGACTCATGATGATGCAATCTTCGAATGCCACATGATGATGATGATGATGAT 1806
 Db 450 GlnAlaAlaGlyIleHisAspAlaIleMetAlaPheProGluGlyTyrArgThnGlnVal 469
 QY 1807 GGGGAACGAGACTCAAGCTTTCAGAGAGAAAGCAAGAAAGTACGATTTGCAAGACC 1866
 Db 470 GlyGluArgGlyLeuLysLeuSerGlyGlyGluLysGlnArgValAlaIleAlaArgThr 489
 QY 1867 ATTTTGAAGACCCCGACGATACCTATGATGAAGCTCTTCAATGCTGATGATGATGAT 1926
 Db 490 IleLeuLysAlaProGlyIleIleLeuLeuAspGluAlaThrSerAlaLeuAspThrSer 509
 QY 1927 ACTGAAGACTATTTCTTGGTGCATGAAGATGTGGTCAAAACAGACACTTCAATTTTC 1986
 Db 510 AsnGluArgAlaIleGlnAlaSerLeuAlaLysValCysAlaAsnArgThrThnIleVal 529
 QY 1987 ATTGCACAGATTTGCAACAGTGTGATGACAGATGAATTCATTTGATCAGGCT 2046
 Db 530 ValAlaHisArgLeuSerThrValAlaAsnAlaAspGlnIleLeuValIleLysAspGly 549
 QY 2047 AAGGTAGCCCAAGTGTGTACCCACCATGCTTGTGCTGATACCTCATAGTATCTATCA 2106
 Db 550 CysIleValGluArgGlyArgHisGluAlaLeuLeuSerArg--GlyGlyValTyrAla 568
 QY 2107 GAATGTGCGATACACAGACAGACCGTGTGACAGAACCATGATTAACCCCAA 2157
 Db 569 AspMetTrpGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 585

Search completed: February 15, 2003, 05:53:54
 Job time : 91 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 01:40:25 ; Search time 3255 Seconds
(without alignments)
11667.716 Million cell updates/sec

Title: AF133659

Perfect score: 2345
Sequence: 1 ATGGCCCTGCTCGCATGCA.....TATTAAAAATCATACATT 2345

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1_0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estdb: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estor: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_man: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2277.6	97.1	2384	11	AF038950 Homo sapi
2	1061.2	45.3	1066	11	U66679 Homo sapien
3	853.8	36.4	1120	13	BM479048 AGENCOURT
4	700.4	29.9	730	14	BM723815 UT-E-EO1-
5	696.6	29.7	787	12	BG722394 602693606
6	685.8	29.2	892	13	BI463667 603207337

7	655.2	27.9	704	9	AU127740
8	653.8	27.9	822	12	BE876927
9	647.8	27.6	840	13	BI549793
10	645	27.5	666	14	BM975299
11	645	27.5	683	12	BG109895
12	633.4	27.0	722	12	BG722908
13	623.2	26.6	817	12	BF310056
14	614	26.2	758	10	BE618331
15	610.4	26.0	627	9	AI769898
16	604.4	25.8	628	14	BM989352
17	587.2	25.0	680	10	BB204009
18	578	24.6	602	14	BG330948
19	576.6	24.6	630	9	AU136079
20	571.6	24.3	618	9	AA668992
21	571	24.3	572	12	BG31020
22	569.8	24.3	811	12	BG069454
23	562	24.0	564	10	AW271773
24	561	23.9	589	10	BE274469
25	561	23.9	611	9	AA642411
26	551.6	23.5	620	9	AU127979
27	548.6	23.4	595	10	BE086841
28	547.2	23.3	700	12	BG122119
29	546.2	23.3	579	10	AA960180
30	545	23.2	546	9	AA305039
31	542	23.1	562	13	BM62767
32	541.8	23.1	559	9	AA912934
33	540.8	23.1	609	9	AA887445
34	530	22.6	546	12	BG31594
35	528.6	22.5	583	13	BM030723
36	525	22.4	549	9	AI016955
37	520.4	22.2	546	10	AW104620
38	518	22.1	534	9	AA923765
39	516.2	22.0	545	9	AA521270
40	511.4	21.8	605	9	AU150336
41	511	21.8	513	9	AI128014
42	507.2	21.6	659	10	BB635137
43	506	21.6	517	9	AA626765
44	505.4	21.6	544	9	AA227515
45	504.2	21.5	589	9	AI678722

ALIGNMENTS

RESULT 1
AF038950
LOCUS AF038950
DEFINITION Homo sapiens ATP binding cassette transporter mRNA, complete cds.
ACCESSION AF038950
VERSION AF038950.1 GI:3329369
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mao, M., Fu, G., Wu, J.-S., Zhang, Q.-H., Zhou, J., Kan, L.-X., Huang, Q.-H., He, K.-L., Gu, B.-W., Han, Z.-G., Shen, Y., Yu, Y.-P., Xu, S.-H., Wang, Y., Chen, S.-J., and Chen, Z.
TITLE Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (14), 8175-8180 (1998)
MEDLINE 98318631
PUBMED 9653160
REFERENCE
AUTHORS 2 (bases 1 to 2384)
Zhang, Q.-H., Ye, M., Wu, X.-Y., Ren, S.-X., Zhao, M., Zhao, G.-J., Fu, G., Shen, Y., Pan, H.-Y., Lu, G., Zhong, M., Xu, X.-R., Han, Z.-G., Zhang, J.-W., Tao, J., Huang, Q.-H., Zhou, J., Hu, G.-X., Gu, J., Chen, S.-J., and Chen, Z.
TITLE Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells
JOURNAL Genome Res. 10 (10), 1546-1560 (2000)

MEDLINE 20499367
PUBMED 11042152
REFERENCE 3 (bases 1 to 2384)
AUTHORS Zhang, Q.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) Rui-Jin Hospital, Shanghai Second Medical University, Shanghai Institute of Hematology, 197 Rui-Jin Road II, Shanghai 200025, P. R. China

FEATURES
source location/Qualifiers
1..2384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq21.3"
/cell_type="CD34+ cell"
/tissue_type="cord blood"
9..2267
/codon_start=1
/product="ATP binding cassette transporter"
/protein_id="AAC39865.1"
/db_xref="GI:3329370"
/translation="MALHMSMRMAAAAFKKRHSALILRLVSVSGSPQMRPH
OLGALGTARAOIPESLSKSTWRLGKNSGFLDAAALQVWPLIEKRTCHHAGS
GLHTDPKGLADVDTKTIKAMLSVWPKDPDLKARVLSLGFVGAKAMNIVPEM
FKYAVDSLQMSGMNLSDAPNTVAATAVLIGYVSRAAFAFENVRNAPGVKA
EONSIRIAKNVFLHNLDFHLSRGTGALSKAIDRTGRGISFVLSLVNPLPHV
EMLVLSGVLYKCCALLNLGILGTATFAVTAFTWRTRFLEIDADNAGNAI
DSLNTYEVKYFNERRVYRVDGFLKTYETASLKSSTLLMLNPGSALFVGLTALI
MYLASQIVACTLTVGDVWVNGHLFQSLPLNPLGTYRRTROALIDMTLFLVLY
DTQIKDKWASPLQTPQATVAEDNVFETIEGQKVLISFVDPKAKVAITVGS
SGKSTIVRLFRFVPEPKGSYLAGQNIQVLSLRVAVVDADALFNTIYNL
LGNISASPEVYVAKLAGLHDAILRMPHYDIOVGRGKLSGGEORVARIALL
KDPVILIDEATSSIDSTIEETILGAMDVVKHRSIPARLSIVVDADETIYLDG
KVAERGTHGLANPHSISYSEMHVQSSRVQNHNPKEAKENISKEEKXQBEI
VMSYKGCNCS"

BASE COUNT 719 a 467 c 567 g 631 t
ORIGIN

Query Match 97.1%; Score 2277.6; DB 11; Length 2384;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2318; Conservative 0; Mismatches 24; Indels 6; Gaps 2;

QY 1 ATGGCGCTGCTGCGATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAAG 60
DB 9 ATGGCGCTGCTGCGATGATCTTGGCGCTGGCGCGCGCGCTTTCGAAAAG 68
QY 61 CGCCGCGCATCCGCGATTTGATCCGCGCTTATGTCCTGTTAGGCGCTCAGTCCGCG 120
DB 69 CGCCGCGCATCCGCGATTTGATCCGCGCTTATGTCCTGTTAGGCGCTCAGTCCGCG 128
QY 121 TGGAGGCCATCACTCGCGCTTGGGAAACGCTCGAGCCATCAAGATTCAGAGTCA 180
DB 129 TGGAGGCCATCACTCGCGCTTGGGAAACGCTCGAGCCATCAAGATTCAGAGTCA 188
QY 181 TTTAAAAGTATCAATGGCAGAGATTGGGAAAAGGCAATTCAGACAGTCTTAAATGCT 240
DB 189 TTTAAAAGTATCAATGGCAGAGATTGGGAAAAGGCAATTCAGACAGTCTTAAATGCT 248
QY 241 GCAAAGGCTCTCGAGTATGGCCATCATGAGAAAAGAGACATGTTGGCATGCTCATGCA 300
DB 249 GCAAAGGCTCTCGAGTATGGCCATCATGAGAAAAGAGACATGTTGGCATGCTCATGCA 308
QY 301 GGAGAGGAGCTCCACACAGACCCAAAGAGGTTAAAAGATGTTGATCTCGAAAATTC 360
DB 309 GGAGAGGAGCTCCACACAGACCCAAAGAGGTTAAAAGATGTTGATCTCGAAAATTC 368
QY 361 ATAAAAGCAATGCTTCTTATGTGTGCCCAAAGACAGGCGAGATCTACAGACTAGAGTT 420
DB 369 ATAAAAGCAATGCTTCTTATGTGTGCCCAAAGACAGGCGAGATCTACAGACTAGAGTT 428
QY 421 GCCATTGGCTGGGATTTTGGGTGTGTGCAAAAGGCATATAATTTGTTCCCTTCATG 480
DB 429 CCCATTGGCTGGGATTTTGGGTGTGTGCAAAAGGCATATAATTTGTTCCCTTCATG 488

QY 481 TTTAAATATGCTGTAGACAGCGCTCAACCGATGTGCGGAAAATGCTGAACTGATGAT 540
DB 489 TTTAAATATGCTGTAGACAGCGCTCAACCGATGTGCGGAAAATGCTGAACTGATGAT 548
QY 541 GCACCAATATACAGTTGACACCATGACCAACAGCACTTCTGATTTGGCTATGCTATCAAGA 600
DB 549 GCACCAATATACAGTTGACACCATGACCAACAGCACTTCTGATTTGGCTATGCTATCAAGA 608
QY 601 GCTGAGCTGCTTTTAAAGAAAGTTGAAATGCAATGATTTGGCAAGTACCCAGAAAT 660
DB 609 GCTGAGCTGCTTTTAAAGAAAGTTGAAATGCAATGATTTGGCAAGTACCCAGAAAT 668
QY 661 TCAATCCGAAGAAATAGCCAAAATGCTTTCTCCATCTTCAACCTGAGCTGAGTTT 720
DB 669 TCAATCCGAAGAAATAGCCAAAATGCTTTCTCCATCTTCAACCTGAGCTGAGTTT 728
QY 721 CACCTGAGCAGACAGACGGAGCTTTATCTAAGCTATTTGACAGAGAACAGGGGTATC 780
DB 729 CACCTGAGCAGACAGACGGAGCTTTATCTAAGCTATTTGACAGAGAACAGGGGTATC 788
QY 781 AGTTTGTCCGAGGCTTGTGTATTTAATCTCTCCCATCATGTTGAAGTATG--- 837
DB 789 AGTTTGTCCGAGGCTTGTGTATTTAATCTCTCCCATCATGTTGAAGTATGCTT 848
QY 838 CTGTGACGTGCTTTGTATTAACAAATGCGGCGCCAGTTGCTTTGTAACCTTGA 897
DB 849 CTGTGACGTGCTTTGTATTAACAAATGCGGCGCCAGTT---GCTTGTACCTTGA 905
QY 898 ACACTTGTATACACACGATTCACAGTTCAGCTCAGCAGGTGAGAACATGATTTGA 957
DB 906 ACACTTGTATACACACGATTCACAGTTCAGCTCAGCAGGTGAGAACATGATTTGA 965
QY 958 ATGAAATGAACAAAGCAGATTAATGATGACAGTATGCTGATAGACTCATGCTGAT 1017
DB 966 ATGAAATGAACAAAGCAGATTAATGATGACAGTATGCTGATAGACTCATGCTGAT 1025
QY 1018 TATGAACTGTGAAGTATTTTAATATGAAGATATGAAGACAGAGATATGATGATTT 1077
DB 1026 TATGAACTGTGAAGTATTTTAATATGAAGATATGAAGACAGAGATATGATGATTT 1085
QY 1078 TTGAAGACGTATGACGCTTCACTTGAAGACCTTCACTTGGCTATGCTGAACCTTT 1137
DB 1086 TTGAAGACGTATGACGCTTCACTTGAAGACCTTCACTTGGCTATGCTGAACCTTT 1145
QY 1138 GGTCAAGTGTATTTTACAGTGTGTTTACAGCTAATAGTGTGCGCAGTCAAGGA 1197
DB 1146 GGTCAAGTGTATTTTACAGTGTGTTTACAGCTAATAGTGTGCGCAGTCAAGGA 1205
QY 1198 ATGTGCGAGGTACCTTACTGTTGAGATCTAGTAATGTGAATGCACTCTTTTTCAG 1257
DB 1206 ATGTGCGAGGTACCTTACTGTTGAGATCTAGTAATGTGAATGCACTCTTTTTCAG 1265
QY 1258 CTTTCATTAACCCCTGAACCTTTCTGGGAACTGTATATAGAGACTAGACAAGCACTCTA 1317
DB 1266 CTTTCATTAACCCCTGAACCTTTCTGGGAACTGTATATAGAGACTAGACAAGCACTCTA 1325
QY 1318 GATATGAACACCTGTTTACTCTACTCAAGGTAGACCCAAATTAAGAACAAGTATG 1377
DB 1326 GATATGAACACCTGTTTACTCTACTCAAGGTAGACCCAAATTAAGAACAAGTATG 1385
QY 1378 GCATCTCCCTTTCAGATCACACACAGACAGCTACCGTGGCTTGTGATATGTCATTTT 1437
DB 1386 GCATCTCCCTTTCAGATCACACACAGACAGCTACCGTGGCTTGTGATATGTCATTTT 1445
QY 1438 GAATACATTTAGGGCCAGAAAGTCTTATGTAATATCTTTGAAGTCCCTGACGAAAG 1497
DB 1446 GAATACATTTAGGGCCAGAAAGTCTTATGTAATATCTTTGAAGTCCCTGACGAAAG 1505
QY 1498 AAGTGGCATTGTAGAGAGTGGGTGAGGGAAGGCAAAATGAGGCTATTAATTT 1557
DB 1506 AAGTGGCATTGTAGAGAGTGGGTGAGGGAAGGCAAAATGAGGCTATTAATTT 1565

QY 1558 CGCTTCTATGAGCTCAAAAGGGTACATTTATCTTGTCTGCTCAAAATATACAAAGATGTG 1617
 Db 1566 CGCTTCTATGAGCTCAAAAGGGTACATTTATCTTGTCTGCTCAAAATATACAAAGATGTG 1625
 QY 1618 AGCCGGAAGACCTTGGAGGGAGTGGAGTGGTACCTAGAGATGCTGCTCTTCAT 1677
 Db 1626 AGCCGGAAGACCTTGGAGGGAGTGGAGTGGTACCTAGAGATGCTGCTCTTCAT 1685
 QY 1678 AATACTATTATTACAACTCTTATATGAAACATACAGTCTTCACTGAGAGATGTAT 1737
 Db 1686 AATACTATTATTACAACTCTTATATGAAACATACAGTCTTCACTGAGAGATGTAT 1745
 QY 1738 GCAAGTGGCAAAATTAAGTGGACCTTCAATGATCAATCTTCAATGCAATGATGAC 1797
 Db 1746 GCAAGTGGCAAAATTAAGTGGACCTTCAATGATCAATCTTCAATGCAATGATGAC 1805
 QY 1798 ACCCAAGTGGGGAAGGAGCTCAAGCTTCAAGAGGAGAAAGCAAAAGTGGCAATT 1857
 Db 1806 ACCCAAGTGGGGAAGGAGCTCAAGCTTCAAGAGGAGAAAGCAAAAGTGGCAATT 1865
 QY 1858 GCAAGAGCCATTTTGAAGGAGCCGCCAGTCACTACTATGATGAGAGTCTTCACTGTTA 1917
 Db 1866 GCAAGAGCCATTTTGAAGGAGCCGCCAGTCACTACTATGATGAGAGTCTTCACTGTTA 1925
 QY 1918 GATTGATTAAGTGAAGACTATTTCTGTGTCATGAAAGATGTGTCAACACAGAACT 1977
 Db 1926 GATTGATTAAGTGAAGACTATTTCTGTGTCATGAAAGATGTGTCAACACAGAACT 1985
 QY 1978 TCTATTTCATGACACAGATTTGTCAACAGTGGTGTATGAGATGAAATCATTTGTCTTG 2037
 Db 1986 TCTATTTCATGACACAGATTTGTCAACAGTGGTGTATGAGATGAAATCATTTGTCTTG 2045
 QY 2038 GATCAGGGTAAAGTGAAGCGGAGCTGTATCCAGCAGTGGTGTCTGTTAACTTCATAGT 2097
 Db 2046 GATCAGGGTAAAGTGAAGCGGAGCTGTATCCAGCAGTGGTGTCTGTTAACTTCATAGT 2105
 QY 2098 ATCTATTCAAAATGTGGCATACACAGAGCCGTGTGCAAAACATGATMAACCCCAA 2157
 Db 2106 ATCTATTCAAAATGTGGCATACACAGAGCCGTGTGCAAAACATGATMAACCCCAA 2165
 QY 2158 TGGGAACCAAAAGAAAGAAATATATCCAAAGAGAGAGAAAGAAACATACAAAGAA 2217
 Db 2166 TGGGAACCAAAAGAAAGAAATATATCCAAAGAGAGAGAAAGAAACATACAAAGAA 2225
 QY 2218 ATTGTCAATAGTGAAGAGCTGTGAAACCTGTCGTAACTGACATAGACATTTTC 2277
 Db 2226 ATTGTCAATAGTGAAGAGCTGTGAAACCTGTCGTAACTGACATAGACATTTTC 2285
 QY 2278 TTTTGTGTGTGTGTGACATATATTGCACTGAGCAGAAATGTTTATTTAAAAAAT 2317
 Db 2286 TTTTGTGTGTGTGTGACATATATTGCACTGAGCAGAAATGTTTATTTAAAAAAT 2345
 QY 2338 CATACATT 2345
 Db 2346 CATACATT 2353
 RESULT 2
 LOCUS U66679 1066 bp mRNA linear HTC 23-JUL-2001
 DEFINITION Homo sapiens clone EST140535 mRNA sequence.
 ACCESSION U66679
 VERSION U66679.1 GI:1906564
 KEYWORDS HTC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1066)
 AUTHORS Allikmets,R., Gerrard,B., Hutchinson,A. and Dean,M.
 TITLE Characterization of the human ABC superfamily: isolation and
 mapping of 21 new genes using the expressed sequence tags database
 JOURNAL Hum. Mol. Genet. 5 (10), 1649-1655 (1996)

MEDLINE 97049974
 PUBMED 8894702
 REFERENCE 2 (bases 1 to 1066)
 AUTHORS Allikmets,R., Gerrard,B. and Dean,M.
 TITLE Direct Submission
 JOURNAL Submitted (12-AUG-1996) Human Genetics Section, National Cancer
 Institute, NCI-FCRDC, Frederick, MD 21702, USA
 FEATURES
 source 1. 1066
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="EST140535"
 /note="similar to ATP-binding cassette transporter"
 BASE COUNT 342 a 201 c 242 g 281 t
 ORIGIN
 Query Match 45.3%; Score 1061.2; DB 11; Length 1066;
 Best Local Similarity 99.7%; Pred. No. 1,6e-223;
 Matches 1063; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1280 TGGGAAGTGTATATAGAGAGACTAGACACACTCATAGATATGAAACACCTTGTACTC 1339
 Db 1 TGGGAAGTGTATATAGAGAGACTAGACACACTCATAGATATGAAACACCTTGTACTC 60
 QY 1340 TACTCAAGTAGACACCCCAATTAAAGACAAAGTAGATGCGATCTCCCTTCAGATCAC 1399
 Db 61 TACTCAAGTAGACACCCCAATTAAAGACAAAGTAGATGCGATCTCCCTTCAGATCAC 120
 QY 1400 CACAGACAGTACCGTGGCTTGTATATGTGATTTGAAATACATTAGAGGCCAGAAAG 1459
 Db 121 CACAGACAGTACCGTGGCTTGTATATGTGATTTGAAATACATTAGAGGCCAGAAAG 180
 QY 1460 TCCCTTAGGAATATCTCTTGAAGTCCCTGACAGAAAGAAAGTGGCATTGTAGAGATA 1519
 Db 181 TCCCTTAGGAATATCTCTTGAAGTCCCTGACAGAAAGAAAGTGGCATTGTAGAGATA 240
 QY 1520 GTGGGTCAGGAAAGACAAATAGTAGAGCTATATTTCGCTTCTATGAGCCTCAAAAG 1579
 Db 241 GTGGGTCAGGAAAGACAAATAGTAGAGCTATATTTCGCTTCTATGAGCCTCAAAAG 300
 QY 1580 GTAGCATTTATCTTGTGTGCTCAAAATATACAGATGTGAGCCTGGAAGCCTTCGAGGG 1639
 Db 301 GTAGCATTTATCTTGTGTGCTCAAAATATACAGATGTGAGCCTGGAAGCCTTCGAGGG 360
 QY 1640 CAGTGGAGAGTGAACCTCAGAGATGCTGCTCTTCATATATACATTTATTAACAACCTC 1699
 Db 361 CAGTGGAGAGTGAACCTCAGAGATGCTGCTCTTCATATATACATTTATTAACAACCTC 420
 QY 1700 TATATGAAACATCAGTGTCTTCACTGAGAAAGTATGAGCGCAAAATTAAGCTGAC 1759
 Db 421 TATATGAAACATCAGTGTCTTCACTGAGAAAGTATGAGCGCAAAATTAAGCTGAC 480
 QY 1760 TTATGATGCAATTTCTTCAATGCCACATGATATGACACCCAAAGTAGGGAACGAGAC 1819
 Db 481 TTATGATGCAATTTCTTCAATGCCACATGATATGACACCCAAAGTAGGGAACGAGAC 540
 QY 1820 TCAAGCTTTCAGAGAGAGAAAGCAAAAGTAGCAATTTGCAAGGCAATTTAAAGGACC 1879
 Db 541 TCAAGCTTTCAGAGAGAGAAAGCAAAAGTAGCAATTTGCAAGGCAATTTAAAGGACC 600
 QY 1880 CCCCACTCATCTATATGATGAGTACTTCACTGTTAGATGATGATGATGATGATGATGAT 1939
 Db 601 CCCCACTCATCTATATGATGAGTACTTCACTGTTAGATGATGATGATGATGATGATGAT 660
 QY 1940 TTCTTGTGTCATGAAAGATGTGTCAACACAGAACTTATTTTCATTTGCAACAGAT 1999
 Db 661 TTCTTGTGTCATGAAAGATGTGTCAACACAGAACTTATTTTCATTTGCAACAGAT 720
 QY 2000 TGTCAACAGTGTGTTGATGAGATGAAATCATTTGTGATGAGGGTAGGGTAGGGTAGGG 2059
 Db 721 TGTCAACAGTGTGTTGATGAGATGAAATCATTTGTGATGAGGGTAGGGTAGGGTAGGG 780
 QY 2060 GTGTGACCCACCATGTTGCTTGTCTTAACCTCATAGTATCTATGAAATGTGCGATA 2119

Db 73 CCAGATTTCATCCGAGATAGCCAAATGTCTTTCTCCATCTTCACACCCGGATCT 132

TITLE Normalization and subtraction: two approaches to facilitate ge

JOURNAL
MEDLINE
COMMENT

discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
Source

Location/Qualifiers
1..730
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EO1-alk-o-22-0-UI"
/clone_1lb="UI-E-EO1"
/cissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-E-EO1 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I and cloned directionally into pRTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (dt)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 213 a 142 c 163 g 211 t 1 others

ORIGIN

Query Match 29.9%; Score 700.4; DB 14; Length 730;

Best Local Similarity 98.2%; Pred. No. 5.7e-144;

Matches 718; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 997 GCTATAGACTCACTGCTGAATTAAGAACTGTGAAGTATTATTAATGAAGATATGAA 1056

Db 1 GCTATAGACTCACTGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60

QY 1057 GCACAGATATGATGATGATTTTGAAGACGTATGAGACTGCTTCATTGAAAGTACTCT 1116

Db 61 GCACAGATATGATGATGATTTTGAAGACGTATGAGACTGCTTCATTGAAAGTACTCT 120

QY 1117 ACTCTGCTATGCTGAATTTGTCAAGTGCATTTTTCAGTGTCCGTTTAAACAGCTATA 1176

Db 121 ACTCTGCTATGCTGAATTTGTCAAGTGCATTTTTCAGTGTCCGTTTAAACAGCTATA 180

QY 1177 ATGTGCTCCCACTGAGGAATTTGTGACAGTACCTTCTGTGTGAGATCTAGTAATG 1236

Db 181 ATGTGCTCCCACTGAGGAATTTGTGACAGTACCTTCTGTGTGAGATCTAGTAATG 240

QY 1237 GTGAATGAGCTGCTTTTCACTTCAATCCCTGAACCTTTCGGAACTGTATATAGA 1296

Db 241 GTGAATGAGCTGCTTTTCACTTCAATCCCTGAACCTTTCGGAACTGTATATAGA 300

QY 1297 GAGACTGACAGACCTATAGATATGACACCTTGTCTACTACAGGTAGACACC 1356

Db 301 GAGACTGACAGACCTATAGATATGACACCTTGTCTACTACAGGTAGACACC 360

QY 1357 CAATTAAGACAAAGTATGATGATCTCCCTTCAGATCACCACAGACGCTACCGTG 1416

Db 361 CAATTAAGACAAAGTATGATGATGATCTCCCTTCAGATCACCACAGACGCTACCGTG 420

QY 1417 GCCTTGATATGATGATGATTTTGAATCATTTAGAGGCGCAAAATCTCTATGAAATTC 1476

Db 421 GCCTTGATATGATGATGATTTTGAATCATTTAGAGGCGCGCAAAATCTCTATGAAATTC 480

QY 1477 TTGAAGTCCCTGACGAAAGAAAGTGGCCATTTGATGAGATGATGATGATGATGATG 1536

Db 481 TTGAAGTCCCTGACGAAAGAAAGTGGCCATTTGATGAGATGATGATGATGATGATG 540

QY 1537 ACAATATGATGATGATGATTTTGTCTTATAGACCTTCAAAAGGATGATGATGATG 1596

Db 541 ACAATATGATGATGATGATTTTGTCTTATAGACCTTCAAAAGGATGATGATGATGATG 600

QY 1597 GGTCAAAATATACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1656

Db 601 GGTCAAAATATACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 659

QY 1657 CAGATGCTGCTCTCTCCATATATATATATATATATATATATATATATATATATATAT 1716

Db 660 CAGATGCTGCTCTCTCTCCATATATATATATATATATATATATATATATATATATAT 719

QY 1717 GCTTCACTGA 1727

Db 720 GCTTCACTGA 730

RESULT 5
BG722394 787 bp mRNA linear EST 08-MAY-2001
LOCUS 602693666F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4825873 5',
DEFINITION mRNA sequence.
ACCESSION BG722394
VERSION BG722394.1 GI:14001581
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 787)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10739 row: 1 column: 02
High quality sequence stop: 771.
Location/Qualifiers
1..787
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4825873"
/clone_1lb="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT	216 a	175 c	211 g	185 t
ORIGIN				
Query Match	29.7%;	Score 696.6;	DB 12;	Length 787;
Best Local Similarity	97.4%;	Pred. No. 3.9e-143;		
Matches 751;	Conservative 0;	Mismatches 14;	Indels 6;	Gaps 4;
Oy	1	ATGGCGCTGCTCGCATCATCTTGGGCGCTGGGCGCGCGCGCTGCTTTGCAAAAG	60	
Db	17	ATGGCGCTGCTCGCATCATCTTGGGCGCTGGGCGCGCGCGCGCTTTCAAAAAG	76	
Oy	61	CGCGGCACTCCGCGATTCGATCCGCGCTTTAGTCTCTTTAGCGGCTCAGATCCGAG	120	
Db	77	CGCGGCACTCCGCGATTCGATCCGCGCTTTAGTCTCTTTAGCGGCTCAGATCCGAG	136	
Oy	121	TGGAGGCCAATCAACTCGGCGCTTTGGGAAACCGCTCCAGGCTAC---CAGATTCCAGAG	177	
Db	137	TGGAGGCCAATCAACTCGGCGCTTTGGGAAACCGCTCCAGGCTACAGAGATTCAGAG	196	
Oy	178	TCATTAATAAGATGATGATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTAAGAT	237	
Db	197	TCATTAATAAGATGATGATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTAAGAT	256	
Oy	238	GCTGCAAAAGGCTCTCCAGGATATGCGCACTGATGAGAAAGAGAGACATGTTGCGATGCTAT	297	
Db	257	GCTGCAAAAGGCTCTCCAGGATATGCGCACTGATGAGAAAGAGAGACATGTTGCGATGCTAT	316	
Oy	298	GCAGAGAGAGAGCTCCACACAGACCCCAAAAGAGGTTTAAAGATGTTGATCTCGGAAA	357	
Db	317	GCAGAGAGAGAGACTCCACACAGACCCCAAAAGAGGTTTAAAGATGTTGATCTCGGAAA	376	
Oy	358	ATCATTAATAAGCAATGCTTTCTTATGTGTGTGCCCAAAAGACAGGCCAGATCTACAGAGCTAGA	417	
Db	377	ATCATTAATAAGCAATGCTTTCTTATGTGTGTGCCCAAAAGACAGGCCAGATCTACAGAGCTAGA	436	
Oy	418	GTTGCCATTTGCTGGGATTTTGGGTGTGCAAAAGGCCATGATATTTGTGTTCCCTTC	477	
Db	437	GTTGCCATTTGCTGGGATTTTGGGTGTGCAAAAGGCCATGATATTTGTGTTCCCTTC	495	
Oy	478	ATGTTTAATAATGCTGTGAGACAGCTCAACAGATGTGGGAAACATGCTGAACCTGAGT	537	
Db	496	ATGTTTAATAATGCTGTGAGACAG-CTCAACAGATGTGGGAAACATGCTGAACCTGAGT	554	
Oy	538	GATGACCAAAATACAGTTGCAACCATGAGCAAGAGTCTGATGGCTATGCTGATATCA	597	
Db	555	GATGACCAAAATACAGTTGCAACCATGAGCAAGAGTCTGATGGCTATGCTGATATCA	614	
Oy	598	AGAGCTGAGCTGCTTTTAAAGAGTTGCAAAATGCAATTTGGCAAGGTAGCCGAG	657	
Db	615	AGAGCTGAGCTGCTTTTAAAGAGTTGCAAAATGCAATTTGGCAAGGTAGCCGAG	674	
Oy	658	AATTCAATCCGAAGAATAAGCAAAATGCTTTCCATCTTCAACACCTGATCTGGGT	717	
Db	675	AATTCAATCCGAAGAATAAGCAAAATGCTTTCTTCAATCTTCAACACCTGATCTGGGT	733	
Oy	718	TTTCACTGAGCAGACAGAGCGGAGCTTATCTAAGGCTATTGACAGAGGA	768	
Db	734	TTTCACTGAGCAGACAGAGCGGAGCTTATCTAAGGCTATTGACAGAGGA	784	
RESULT 6				
LOCUS	B1463667	892 bp	mRNA	linear
DEFINITION	603207337F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5273168 5',			
ACCESSION	B1463667			
VERSION	B1463667.1 GI:15254323			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 892)			

AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L1M1689 row: 1 column: 09 High quality sequence stop: 721. Location/Qualifiers				
FEATURES	1. 892				
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5273168" /clone_1id="NIH_MGC_97" /lab_host="DH10B" /note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site 1: BamI; Site 2: SalI-XhoI (gtcgaag) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."				
BASE COUNT	231 a	201 c	237 g	223 t	
ORIGIN					
Query Match	29.2%;	Score 685.8;	DB 13;	Length 892;	
Best Local Similarity	95.0%;	Pred. No. 9.3e-141;			
Matches 829;	Conservative 0;	Mismatches 27;	Indels 17;	Gaps 11;	
1	ATGGCGCTGCTCGCATCATCTTGGGCGCTGGGCGCGCGCTGCTTTGCAAAAG	60			
17	ATGGCGCTGCTCGCATCATCTTGGGCGCTGGGCGCGCGCTGCTTTGCAAAAG	76			
61	CGCGCGCACTCCGCGATTCGATCCGCGCTTTAGTCTCTGTTAGCGGCTCAGTCCGAG	120			
77	CGCGCGCACTCCGCGATTCGATCCGCGCTTTAGTCTCTGTTAGCGGCTCAGTCCGAG	136			
121	TGGAGGCCAATCAACTGGGCGCTTTGGGAAACCGCTGAGCTAC---CAGATTCCAGAG	177			
137	TGGAGGCCAATCAACTGGGCGCTTTGGGAAACCGCTGAGCTACCAAGATTCAGAG	196			
178	TCATTAATAAGATGATGATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTAAGAT	237			
197	TCATTAATAAGATGATGATGAGAGATTTGGGAAAGGCAATTCAGACAGTCTTAAGAT	256			
238	GCTGCAAAAGGCTCTCCAGGATATGCGCACTGATGAGAAAGAGAGATTTGGCATGCTAT	297			
257	GCTGCAAAAGGCTCTCCAGGATATGCGCACTGATGAGAAAGAGAGATTTGGCATGCTAT	316			
298	GCAGAGAGAGAGCTCCACACAGACCCCAAAAGAGGTTTAA--AGATCTGATCTCGGA	355			
317	GCAGAGAGAGAGCTCCACACAGACCCCAAAAGAGGTTTAA--AGATCTGATCTCGGA	376			
356	AAATCATTAATAAGATGCTTTTATGTTGTTGGCCCAAGACAGGCGAGATCTAGAGTA	415			
377	AAATCATTAATAAGATGCTTTTATGTTGTTGGCCCAAGACAGGCGAGATCTAGAGTA	436			
416	GAGTTGCCATTTGCTGGGATTTTGGGTGTGCAAAAGGCCATGATATTTGGTTCCT	475			
437	GAGTTGCCATTTGCTGGGA--TTTGGGTGTGCAAAAGGCCATGATATTTGGTTCCT	495			
476	TCATGTTTAATAATGCTGTAGACAGCTCAACCAATGTGGGAAACATGCTGAACCTGA	535			
496	TCATGTTTAATAATGCTGTAGACAGCTCAACCAATGTGGGAAACATGCTGAACCTGA	555			

QY 536 GTGATGCACCAATACAGTTGCAACATGGCAACAGAGTTTGTGATGGCTATG 595
| | | | |
Db 556 GTGATGCACCAATACAGTTGCAACATGGCAACAGAGTTTGTGATGGCTATG 615
| | | | |
QY 596 CAAGAGCTGGAGCTGCTTTTAAACGAAGTTGCAAGTATTTGGCAAGTAG-CC 654
| | | | |
Db 616 CAAGAGCTGGAGCTGCTTTTAAACGAAGTTGCAAGTATTTGGCAAGTAGCCCC 675
| | | | |
QY 655 CAGATTCATCCAGAGATA-GCCAAAATGTCTT-CTCCATCTTCAACAACCTGG- 709
| | | | |
Db 676 CAGATTCATCCAGAGATAATGCCAAACGTCTTACTCTTCAACAACCTGGCAT 735
| | | | |
QY 710 ATCTGGGTTTACCTGAGCAGACA-GACGGAGCTTATCT-AAGCTATTTGACAGAG 767
| | | | |
Db 736 CTGGGGTTTACCTGAGCAGACACGCGAGCTTATCTAAGGCTATTTGACAGAG 795
| | | | |
QY 768 AACAA--GGGTATCATTTTGTCTGAGTCTTGGTATTTATCTTCTCCATCATG 825
| | | | |
Db 796 GACCAAGCGGTATCATTTTGTCTGAGTCTTGGTATTTATCTTCTCCATCATG 855
| | | | |
QY 826 TTGAGAGTGAAGCTTGTCAAGTGTGTTTGTAT 858
| | | | |
Db 856 -TTGAAATGATGCTTGTCAAGAGGCTTGTATTT 887
| | | | |

RESULT 7

AU127740

LOCUS AU127740 NT2RP2 Homo sapiens cDNA clone NT2RP2001958 5', mRNA

DEFINITION

sequence.

ACCESSION AU127740 GI:10988094

VERSION AU127740

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 704)

Oca.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,

Yamamoto.J., Makamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and

Isogai.T.

HRI human cDNA project

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1. 704

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2RP2001958"

/clone_lib="NT2RP2"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/note="Vector: pME18SFL3; RNA from NT2 neuronal precursor

cells after 2-weeks retinoic acid (RA) induction."

BASE COUNT 190 a 142 c 168 g 198 t

ORIGIN

Query Match 27.9%; Score 655.2; DB 9; Length 704;

Best Local Similarity 97.0%; Pred. No. 5.3e-134;

Matches 665; Conservative 0; Mismatches 19; Indels 2; Gaps 2;

QY 232 TTAGATGCTCAAGGCTCTCCAGTATGCGCACTGATAGAAAAAGAGCATGTTGGCAT 291

Db 1 TTAGATGCTCAAGGCTCTCCAGTATGCGCACTGATAGAAAAAGAGCATGTTGGCAT 60
| | | | |
QY 292 GGTATGTCAGAGAGAGAGACTCCACAGACCCCAAGAGGTTTAAAGATGTTGATCT 351
| | | | |
Db 61 GGTATGTCAGAGAGAGAGACTCCACAGACCCCAAGAGGTTTAAAGATGTTGATCT 120
| | | | |
QY 352 CGGAAATATCAATAAGCAATGCTTTCTTATGTGTGCCCCAAGACAGGCCAGATCTACGA 411
| | | | |
Db 121 CGGAAATATCAATAAGCAATGCTTTCTTATGTGTGCCCCAAGACAGGCCAGATCTACGA 180
| | | | |
QY 412 GCTAGAGTTGCCATTTTCTGCTGGGATTTTGGGTGTGTGCAAGGCCATGATATTGTGCTT 471
| | | | |
Db 181 GCTAGAGTTGCCATTTTCTGCTGGGATTTTGGGTGTGTGCAAGGCCATGATATTGTGCTT 240
| | | | |
QY 472 CCTTATGTTTAAATATGCTGTGAGACAGCTTCAACAGATGTGCGGAAACATGCTGAAC 521
| | | | |
Db 241 CCTTATGTTTAAATATGCTGTGAGACAGCTTCAACAGATGTGCGGAAACATGCTGAAC 300
| | | | |
QY 532 CTGAGTATGACCAATAATACAGTTGCAACCATGGCAACAGCAGTTGCTATGCTATGCT 591
| | | | |
Db 301 CTGAGTATGACCAATAATACAGTTGCAACCATGGCAACAGCAGTTGCTATGCTATGCT 360
| | | | |
QY 592 GTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTGCAAGTATGCTATGCTATGCTATGCT 651
| | | | |
Db 361 GTATCAAGAGCTGAGCTGCTTTTAAACGAAGTTGCAAGTATGCTATGCTATGCTATGCT 420
| | | | |
QY 652 GCCCAGATTCATCCGAAATAGCCAAATATGCTTCTCATCTTCAACACCTGGAT 711
| | | | |
Db 421 GCCCAGATTCATCCGAAATAGCCAAATATGCTTCTCATCTTCAACACCTGGAT 480
| | | | |
QY 712 CTGGGTTTCACTGAGCAGACAGACGAGAGCTTTATCTAAGGCTATTTGACAGAGAGACA 771
| | | | |
Db 481 CTGGGTTTCACTGAGCAGACAGACGAGAGCTTTATCTAAGGCTATTTGACAGAGAGACA 539
| | | | |
QY 772 AGGGGTATCAGTTTGTCTGAGTCTTGTGATTTATCTTCTCCATCATGTTTGA 831
| | | | |
Db 540 AGGGGTATCAGTTTGTCTGAGTCTTGTGATTTATCTTCTCCATCATGTTTGA 598
| | | | |
QY 832 GTGATGCTGTGAGCTGCTTTTGTATTAACAAATGGGCGCCAGTTGCTTGTGTAAC 891
| | | | |
Db 599 GTGATGCTGTGAGCTGCTTTTGTATTAACAAATGGGCGCCAGTTGCTTGTGTAAC 658
| | | | |
QY 892 CTGGAACACTTGTGATACATACAGCATTCACAGTTGCTGACATC 937
| | | | |
Db 659 CTTGGAACACTTGTGATACATACAGCATTCACAGTTGCTGACATC 704
| | | | |

RESULT 8
BE876927 822 bp mRNA linear EST 20-OCT-2000
LOCUS BE876927
DEFINITION 601484702P1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3887161 5',
mRNA sequence.
ACCESSION BE876927
VERSION BE876927 GI:10325690
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 822)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov

Oy	300	AGGAGAGGACCTCCACACAGACC	CAAAGAAGGTTTAAAGATGTGA-TACTCGGAAAT	359
Dd	315	AGGAGAGGAGACTCCACACAGACC	CAAAAAGGGTTTAAAGATGTGA-TACTCGGAAAT	374
Oy	360	CATTAAGCAATGCCTTCTTATGTG	TGGGCCCAAAGACAGCAGATCTACGAGTAGAGT	419
Dd	375	CATTAAAGCAATGCCTTCTTATGTG	TGGGCCCAAAGACAGCAGATCTACGAGTAGAGT	434
Oy	420	TGCCATTTGCTGGAGATTTTTGGTG	TGTCAAAGGCATGAATTAATGTGTTCCCTTCAT	479
Dd	435	TGCCATTTGCTGGAGATTTTTGGTG	TGTCAAAGGCATGAATTAATGTGTTCCCTTCAT	494
Oy	480	GTTTAAATAATGTGTAGACAGCCT	CAACCAATGTTCGGAAACATGCTGAACCTAGTGA	539
Dd	495	GTTTAAATAATGTGTAGACAGCCT	CAACCAATGTTCGGAAACATGCTGAACCTAGTGA	554
Oy	540	TGCACCAAAATACAGTTGCACAAC	CATGAGCACAGATTCGATTGGCTAATGGTATACAG	599
Dd	555	TGCACCAAAATACAGTTGCACAAC	CATGAGCACAGATTCGATTGGCTAATGGTATACAG	613
Oy	600	AGCTGGAGCTGCTTTTTTTAACGAA	ATTGCAATTTGGCAAGTAGCCACAGNA	659
Dd	614	AGCTGGAGCTGCTTTTTTTAACGAA	ATTGCAATTTGGCAAGTAGCCACAGNA	673
Oy	660	TTTCATTCGCAAAATATGCCAAAA	AATGCTTTTCCCATCTTCCAACTCGATCTGGG--T	717
Dd	674	TTTCATTCGCAAAATATATAGCC-AAA	AATGCTTTTCCCATCTTCCAACTCGATCTGGGTTT	732
Oy	718	TTTCACCTGAGCAGACA	734	
Dd	733	TTTCACCTGAGCCGACA	749	
RESULT 10				
LOCUS	BM975299/c			
DEFINITION	BM975299	666 bp	mRNA	linear EST 21-MAR-2002
ACCESSION	U1-CF-ECL1-ack-e-24-0-U1.s1		U1-CF-ECL1 Homo sapiens	cDNA clone
VERSION	BM975299			
KEYWORDS	BM975299.1 GI:19592890			
SOURCE	EST.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL MEDLINE COMMENT	Bonaldo M.F., Lennon G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9) , 791-806 (1996) 97044477 Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel.: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). Seq primer: M13 FORWARD POLYA=yes. Location/Qualifiers 1..666 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="U1-CF-ECL1-ack-e-24-0-U1" /clone_1lib="U1-CF-ECL1" /tissue_type="Lung"			

```

/dev stage="Adult and Fetal"
//lab_host="DH10B (Life Technologies) (T1 phage resistant)"
//note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-ECl is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGGCTTAC.
TAG LIB=UI-CF-ECl
TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 360-363
TAG SEQ=AAGTCTTAC"
169 a 145 c 118 g 234 t
BASE COUNT
ORIGIN

```

Query Match	27.5%	Score 645	DB 14	Length 666
Best Local Similarity	100.0%	Pred. No. 9,4e-132		
Matches 645	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1701	ATATGAAACATCGAGTCTTCACCTGAGGAGAGTGTATGCAAGTGGCAAAATTAGCTGACT	1760	
Db	666	ATATGAAACATCGAGTCTTCACCTGAGGAGAGTGTATGCAAGTGGCAAAATTAGCTGACT	607	
QY	1761	TCATGATGCAATTCTTTGGAATGCGACATGGAATGACACCCAAAGTGGGAAACGAGACT	1820	
Db	606	TCAATGATGCAATTCTTTGGAATGCGACATGGAATGACACCCAAAGTGGGAAACGAGACT	547	
QY	1821	CAAGCTTTACGAGAGAGAAAAAGCAAAAGTAGCAATTGCAAGACCAATTTTGAAGACC	1880	
Db	546	CAAGCTTTACGAGAGAGAAAAAGCAAAAGTAGCAATTGCAAGACCAATTTTGAAGACC	487	
QY	1881	CCCAAGTCATCTCTATGATGAGACTCTCATGCTTGAATTCGATTACTGAGAGACTAT	1940	
Db	486	CCCAAGTCATCTCTATGATGAGACTCTCATGCTTGAATTCGATTACTGAGAGACTAT	427	
QY	1941	TCTTGTGCGCATGAAAGGTGGTGCAAACACAGAACTTCAATTTTCATGACACAGATT	2000	
Db	426	TCTTGTGCGCATGAAAGGTGGTGCAAACACAGAACTTCAATTTTCATGACACAGATT	367	
QY	2001	GTCAACAGTGGTTGATGACAGTGAATCATTTGCTTGGATCAGGGTAAAGGTAGCCGAAAG	2060	
Db	366	GTCAACAGTGGTTGATGACAGTGAATCATTTGCTTGGATCAGGGTAAAGGTAGCCGAAAG	307	
QY	2061	TGATACCCACCATGCTTGCCTGCTAACCTCCATAGATATCATTCACAAAATGTCGATAC	2120	
Db	306	TGATACCCACCATGCTTGCCTGCTAACCTCCATAGATATCATTCACAAAATGTCGATAC	247	
QY	2121	ACAGAGACGCGTGTGACGAACCATGATTAACCCCAATGGGAGCAAGAAAGAAATAT	2180	
Db	246	ACAGAGACGCGTGTGACGAACCATGATTAACCCCAATGGGAGCAAGAAAGAAATAT	187	
QY	2181	ATCCAAAGAGAGAGAAAGAAAGAACTACAGAAAGAAATGTCAATGTGTGAAGGCTG	2240	
Db	186	ATCCAAAGAGAGAGAGAAAGAAAGAACTACAGAAAGAAATGTCAATGTGTGAAGGCTG	127	
QY	2241	TGAGAACTGTTGCGTAAAGTACATTAAGCATTTTCTTTTTTGTGTTTGGACTACA	2300	
Db	126	TGAGAACTGTTGCGTAAAGTACATTAAGCATTTTCTTTTTTGTGTTTGGACTACA	67	
QY	2301	TATTTGACATGAGACAGAAATGTTTTATTAATAAAATCATATCATT	2345	
Db	66	TATTTGACATGAGACAGAAATGTTTTATTAATAAAATCATATCATT	22	

```
RESULT 11
BG109895          683 bp      mRNA      linear      EST 30-JAN-2001
LOCUS             602279565F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4367128 5',
DEFINITION        mRNA sequence.
ACCESSION         BG109895
VERSION           BG109895.1 GI:12603401
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 683)
AUTHORS           NIH-MGC http://mhc.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  Plate: LHAM10019 row: a column: 17
                  High quality sequence stop: 681.
FEATURES
  source
    1..683
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:4367128"
       /clone_1lb="NIH_MGC_86"
       /tissue_type="osteosarcoma, cell line"
       /lab_host="DH10B (phage-resistant)"
       /note="Organ: bone; Vector: pCMV-SPORT6, Site_1: NotI;
       Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
       Average insert size 1.533 kb. Library enriched for
       full-length clones and constructed by Life Technologies.
       Note: this is a NIH_MGC library."
BASE COUNT      232 a 125 c 145 g 180 t 1 others
ORIGIN
Query Match      27.5%; Score 645; DB 12; Length 683;
Best Local Similarity 99.8%; Pred. No. 9,4e-132;
Matches 656; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1681 ACTATTATTTACACCTCTTATATGAAACATCACTGCTTCACTGAGAGATGATGCA 1740
DB 1 ACTATTATTTACACCTCTTATATGAAACATCACTGCTTCACTGAGAGATGATGCA 60

QY 1741 GTGGGAAAATTGCTGAGACTTATGATGATGCAATTTCTTGAATGCCACATGATATGACACC 1800
DB 61 GTGGGAAAATTGCTGAGACTTATGATGATGCAATTTCTTGAATGCCACATGATATGACACC 120

QY 1801 CAAGTAGGGGAACGAGACTCAAGCTTTTCAGAGAGAGAAAACAAAGAGTAGCAATTGCA 1860
DB 121 CAAGTAGGGGAACGAGACTCAAGCTTTTCAGAGAGAGAAAACAAAGAGTAGCAATTGCA 180

QY 1861 AGAGCCATTTTGAAGAGACCCCCAGCTCATATCTCTATGATGAAGACTTCTCATGTTAAGT 1920
DB 181 AGAGCCATTTTGAAGAGACCCCCAGCTCATATCTCTATGATGAAGACTTCTCATGTTAAGT 240

QY 1921 TCGATTACTGAAGAGACTTTTGTGTGTCATGAAGAGATGTGTCAACAGCAACTTCT 1980
DB 241 TCGATTACTGAAGAGACTTTTGTGTGTCATGAAGAGATGTGTCAACAGCAACTTCT 300

QY 1981 ATTTTCATTTGACACAGATTTGTCACAGTGTGTGATGAGATGAATCATTTGTTGAT 2040
DB 301 ATTTTCATTTGACACAGATTTGTCACAGTGTGTGATGAGATGAATCATTTGTTGAT 360

QY 2041 CAGGGTAAGTAGCCGAACGTGTATCCACCATGTTGCTTCAACCTCATAGTATC 2100
|||||
```

```
DB 361 CAGGGTAAGTAGCCGAACGTGTATCCACCACGTTTCTGCTAACCTCATATGATC 420

QY 2101 TATTCAGAAATGTGTCATACACAGACGCCGTGTGAGAACCATGATTAACCCCAATGG 2160
DB 421 TATTCAGAAATGTGTCATACACAGACGCCGTGTGAGAACCATGATTAACCCCAATGG 480

QY 2161 GAAGCAAAAGAAAATAATATTCCTCAAGAGAGAGAAAGAAAGAACTTACAGAAAGAAATT 2220
DB 481 GAAGCAAAAGAAAATAATATTCCTCAAGAGAGAGAAAGAAAGAACTTACAGAAAGAAATT 540

QY 2221 GTCAATAGTGTGAAGAGCTGTGGAAGAACTGTCGTCCTAATGACATATGACATTTTCT 2279
DB 541 GTCAATAGTGTGAAGAGCTGTGGAAGAACTGTCGTCCTAATGACATATGACATTTTCTT 600

QY 2280 TTTTGTGTTTGTGACATACATATTTGACAGAGAAATGTTTATTAATAAAAA 2336
DB 601 TTTTGTGTTTGTGACATACATATTTGACAGAGAAATGTTTATTAATAAAAA 657

RESULT 12
BG722908          722 bp      mRNA      linear      EST 08-MAY-2001
LOCUS             602659410F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4827578 5',
DEFINITION        mRNA sequence.
ACCESSION         BG722908
VERSION           BG722908.1 GI:14002095
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 722)
AUTHORS           NIH-MGC http://mhc.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
                  CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
                  Toshiyuki and Piero Carninci (RIKEN)
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  Plate: LHAM10744 row: c column: 03
                  High quality sequence stop: 710.
FEATURES
  source
    1..722
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:4827578"
       /clone_1lb="NIH_MGC_97"
       /lab_host="DH10B"
       /note="Organ: testis; Vector: pBluescript (modified
       pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (9tc9ag
       ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTVA-3',
       size-selected for average insert size 2.2 kb and
       normalized to Rot 5. This is a primary library enriched
       for full-length clones and constructed using the
       Cap-trapper method (Carninci, in preparation). Library
       constructed by M. Brownstein (NIH/NHGRI, National
       Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT      191 a 167 c 191 g 173 t
ORIGIN
Query Match      27.0%; Score 633.4; DB 12; Length 722;
Best Local Similarity 99.0%; Pred. No. 3,4e-129;
Matches 670; Conservative 0; Mismatches 1; Indels 6; Gaps 3;

QY 1 ATGGCGCTGCTCGGAGTCAATTTCTTGGCGCTGGCGCGCGGCGGCTGTTGAAAAAG 60
DB 44 ATGGCGCTGCTCGGAGTCAATTTCTTGGCGCTGGCGCGCGGCGGCGGCTGTTGAAAAAG 103
|||||
```

QY 61 CGCGGCACTCCGCGATTCTGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAAGTCCGAG 120
Db 104 CGCGGCACTCCGCGATTCTGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAAGTCCGAG 163
QY 121 TGGAGGCAATCACTCGGCGCCTTGGGAAACCGCTGAGCTAC---CAGATTCCAGAG 177
Db 164 TGGAGGCAATCACTCGGCGCCTTGGGAAACCGCTGAGCTACCGAGCAATTCAGAG 223
QY 178 TCATTAATAAGTATCAGATGCGAGAGATTGGGAAAAGCAATTCAGAGCACTTCTTAGAT 237
Db 224 TCATTAATAAGTATCAGATGCGAGAGATTGGGAAAAGCAATTCAGAGCACTTCTTAGAT 283
QY 238 GCTGCAAAAGCTCTTCAGAGTATGCGCACTGATAGAAAAGAGACATGTTGCGATGTCAT 297
Db 284 GCTGCAAAAGCTCTTCAGAGTATGCGCACTGATAGAAAAGAGACATGTTGCGATGTCAT 343
QY 298 GCAGAGAGAGAGACTCCACAGACCCAAAAGAGGTTAAAGATGTTGATACCTCGGAAA 357
Db 344 GCAGAGAGAGAGACTCCACAGACCCAAAAGAGGTTAAAGATGTTGATACCTCGGAAA 403
QY 358 ATCATTAATAAGCAATGCTTTCTTATGTGTGGCCCAAAAGACAGCCAGATCTACGAGCTAGA 417
Db 404 ATCATTAATAAGCAATGCTTTCTTATGTGTGGCCCAAAAGACAGCCAGATCTACGAGCTAGA 463
QY 418 GTTGCCATTTCGCTGGGATTTTGGGTGGTCAAAAGGCGCATGAATATTGTGTTCCCTTC 477
Db 464 GTTGCCATTTCGCTGGG--TTTGGGTGGTCAAAAGGCGCATGAATATTGTGTTCCCTTC 521
QY 478 ATGTTTAAATATGCTGTAGAGAGGCTCAACAGATGTCTGGGAAACATGCTGAACCTTAGT 537
Db 522 ATGTTTAAATATGCTGTAGAGAGGCTCAACAGATGTCTGGGAAACATGCTGAACCTTAGT 581
QY 538 GATGACCAATAATAGATTGCAACCATGCGAACAGAGTTCGATGTGCTATGCTATCA 597
Db 582 GATGACCAATAATAGATTGCAACCATGCGAACAGAGTTCGATGTGCTATGCTATCA 641
QY 598 AGAGCTGAGC-TGCTTTTAAAGAAAGTTGGAATGCAATGCTATTTGGCAGGTAGCCCA 656
Db 642 AGAGCTGAGCTGCTTTTAAAGAAAGTTGGAATGCAATGCTATTTGGCAGGTAGCCCA 701
QY 657 GAATTCATCCGAGAA 673
Db 702 GAATTCATCCGAGAA 718

RESULT 13
BF310056 817 bp mRNA linear EST 21-NOV-2000
LOCUS 601892556F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4138236 5',
DEFINITION mRNA sequence.
ACCESSION BF310056
VERSION BF310056.1 GI:11257555
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph. D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
plate: LLCM046 row: p column: 13
High quality sequence stop: 644.
Location/Qualifiers

source 1. 817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4138236"
/clone_1ib="NIH MGC 17"
/cistug_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pORF7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 207 a 192 c 227 g 191 t
ORIGIN
Query Match 26.6%; Score 623.2; DB 12; Length 817;
Best Local Similarity 96.8%; Pred. No. 5.9e-127;
Matches 722; Conservative 0; Mismatches 13; Indels 11; Gaps 8;
QY 1 ATGGCGCTGCTCGCGATTCGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAGGTCGCGAG 120
Db 3 ATGGCGCTGCTCGCGATTCGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAGGTCGCGAG 62
QY 61 CGCGGCACTCCGCGATTCTGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAGGTCGCGAG 120
Db 63 CGCGGCACTCCGCGATTCTGATCCGGCCTTTAGTCTCTGTTAGCGGCTCAGGTCGCGAG 122
QY 121 TGGAGGCAATCACTCGGCGCCTTGGGAAACCGCTGAGCTTAC---CAGATTCCAGAG 177
Db 123 TGGAGGCAATCACTCGGCGCCTTGGGAAACCGCTGAGCTTACCGAGATTCAGAG 182
QY 178 TCATTAATAAGTATCAGATGCGAGAGATTGGGAAAAGCAATTCAGAGCACTTCTTAGAT 237
Db 183 TCATTAATAAGTATCAGATGCGAGAGATTGGGAAAAGCAATTCAGAGCACTTCTTAGAT 241
QY 238 GCTGCAAAAGCTCTCGAGGATGAGGCACTGATAGAAAAGAGCATGTTGCGATGTCAT 297
Db 242 GCTGCAAAAGCTCTCGAGGATGAGGCACTGATAGAAAAGAGCATGTTGCGATGTCAT 300
QY 298 GCAGAGAGAGAGACTCCACAGACCCAAAAGAGGTTAAAGATGTTGATCTCGGAAA 357
Db 301 GCAGAGAGAGAGACTCCACAGACCCAAAAGAGGTTAAAGATG-TGATATCTCGGAAA 359
QY 358 ATCATTAATAAGCAATGCTTTCTTATGTGTGGCCCAAAAGACAGGCGCATCTACGAGCTAGA 417
Db 360 ATCATTAATAAGCAATGCTTTCTTATGTGTGGCCCAAAAGACAGGCGCATCTACGAGCTAGA 419
QY 418 GTTGCCATTTCGCTGGGATTTTGGGTGGTCAAAAGGCGCATGAATATTGTGTTCCCTTC 477
Db 420 GTTGCCATTTCGCTGGGATTTTGGGTGGTCAAAAGGCGCATGAATATTGTGTTCCCTTC 479
QY 478 ATGTTTAAATATGCTGTAGAGAGGCTCAACAGATGTGCGGAAAACATGCTGAACCTTAGT 537
Db 480 ATGTTTAAATATGCTGTAGAGAGGCTCAACAGATGTGCGGAAAACATGCTGAACCTTAGT 539
QY 538 GATGACCAATAATAGATTGCAACCATGCGAACAGAGTTCGATGTGCTATGCTATCA 597
Db 540 GATGACCAATAATAGATTGCAACCATGCGAACAGAGTTCGATGTGCTATGCTATCA 598
QY 598 AGAGCTGAGAGCTCTTTTAAAGAAAGTTGGAATGCAATGCTATTTGGCAAGGTAGCCAG 657
Db 599 AGAGCTGAGAGCTCTTTTAAAGAAAGTTGGAATGCAATGCTATTTGGC-AGGTAGCCAG 657
QY 658 AATTCAATCCGAAGATAGCAAAAATGCTTTC-TCCATCTTCACAA--CTGGAATCTG 714
Db 658 AATTCAATCCGAAGATAGCAAAAATGCTTTC-TCCATCTTCACAACTCGATCTCG 717
QY 715 GGTTCACCTGAGCAGACAGACGGG 740
Db 718 GGTTCACCTGAGCAGACAGACGGG 743

RESULT 14
LOCUS BE618331 758 bp mRNA linear EST 20-OCT-2000
DEFINITION 601462724F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:386633 5',
mRNA sequence.
ACCESSION BE618331
VERSION BE618331
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 758)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9610 row: 0 column: 06
High quality sequence stop: 648.
Location/Qualifiers
source 1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:386633"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectional. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 214 a 165 c 195 g 184 t
ORIGIN
Query Match 26.2% Score 614; DB 10; Length 758;
Best Local Similarity 94.4%; Pred. No. 6.3e-125;
Matches 714; Conservative 0; Mismatches 30; Indels 12; Gaps 7;
QY 74 CGATTCTGATCCGGCTTGTAGTCTCTGTAGCGGCTCAAGTCCGACAGTGGAGCCACATC 133
Db 1 CGATTCTGATCCGGCTTGTAGTCTCTGTAGCGGCTCAAGTCCGACAGTGGAGCCACATC 60
QY 134 AACTGGGCGCTTTGGGAACCGCTCGAGCTTAC---CAGATTCCAGAGTCATTAAAAAGT 189
Db 61 AACTGGGCGCTTTGGGAACCGCTCGAGCTTACAGGCAAGTCCAGAGTCATTAAAAAGT 120
QY 190 ATCAGATGGCAAGATTTGGAAAAGGCAATTCAGACAGTCTCTTGAATCTGCACAAAGCT 249
Db 121 ATCAGATGGCAAGATTTGGAAAAGGCAATTCAGACAGTCTCTTGAATCTGCACAAAGCT 180
QY 250 CTCGAGTATGGCCACTGATAGAAAAAGAGACATGTTGGCATGTGCATGACAGAGAGA 309
Db 181 CTCGAGTATGGCCACTGATAGAAAAAGAGACATGTTGGCATGTGCATGACAGAGAGA 240
QY 310 CTCGACACAGATCCCAAAAGAGGTTAAAAAGATGTTGATATCTCGAAAAATCATAAAAAGA 369
Db 241 CTCGACACAGATCCCAAAAGAGGTTAAAAAGATGTTGATATCTCGAAAAATCATAAAAAGA 300
QY 370 ATGCTTTCTTATGTGGGGCCCAAGACAGGCGAGATCTTCAGAGCTTGGCCATTTCG 429
Db 301 ATGCTTTCTTATGTGGGGCCCAAGACAGGCGAGATCTTCAGAGCTTGGCCATTTCG 360
QY 430 CTGGGATTTTGGGTGGTGCAAAAGGCCATGAATATTGGTTCCTTCATGTTTAAATAT 489

Db 361 CTGGGATTTTGGGTGGTGCAAAAGGCCATGAATATTGGTTCCTTCATGTTTAAATAT 420
QY 490 GCTGTAGACAGCTTCACCAACGATGTGGGAAACATGCTGAACCTGAGTATGACCAAAAT 549
Db 421 GCTGTAGACAGCTTCACCAACGATGTGGGAAACATGCTGAACCTGAGTATGACCAAAAT 480
QY 550 ACAGTTGCACCATGAGCAACAGCAGTTCTGATTGGCTATGGTATCAAGAGCTGAGCT 609
Db 481 ACAGTTGCACCATGAGCAACAGCAGTTCTGATTGGCTATGGTATCAAGAGCT -GAGCT 539
QY 610 GCTTTTAAACGAATTCGAATTCAGATATTGGCAAGGTAGGCCGAATTCATCCGA 669
Db 540 GCTTTTAAACGAATTCGAATTCAGATATTGGCAAGGTAGGCCGAATTCATCCGA 599
QY 670 -AGAAATGCCAAAATATGTC-TTCTCATCTTCACAACCTGGATCTGGTTTCACTGA 727
Db 600 CAGAAATGCCAAAATATGTCCTTCTCATCTTCACAACCTGGATCTGGTTTCACTTGA 659
QY 728 GCAG---ACAGACGCGAGCTTATCTAAGGCTATTGAC-AGAGAACAGAGGGTA-TCAG 782
Db 660 GCAGAACAGAGCGGCGAAGCTATCTAAGGCTATTGACAGAACAGAGGCGATTTCAG 719
QY 783 TTTTGCTGTAGTCTTTGTTGATTAAATCTTCTTC 818
Db 720 GTTGTCCGAGAGCTCGGAGTCAAGTCTTCTCC 755
RESULT 15
A1769898/c 627 bp mRNA linear EST 21-DEC-1999
LOCUS wj30c01.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2404320 3'
DEFINITION similar to SW:ABC7 HUMAN 075027 ATP-BINDING CASSETTE TRANSPORTER 7
PRECEDSOR ;, mRNA sequence.
ACCESSION A1769898
VERSION A1769898.1 GI:5236407
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 627)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrip/image/image.html
Insert length: 556 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 473.
Location/Qualifiers
source 1..627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2404320"
/clone_lib="NCI-CGAP-Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP Kids was
prepared, and 98 circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subclative
hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fátima Bonaldo.

BASE COUNT 164 a 140 c 111 g 212 t
ORIGIN

Query Match 26.0%; Score 610.4; DB 9; Length 627;
Best Local Similarity 99.0%; Pred. No. 4e-124;
Matches 614; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1726 GAGGAAGTATGCAAGTGGCAAAATTAAGTCTGATGCAATTTCTGAAATGCCA 1785
|||
DB 627 GAGGAAGTATGCAAGTGGCAAAATTAAGTCTGATGCAATTTCTGAAATGCCA 568
|||
QY 1786 CATGATATGACACCCAAAGTGGGAAACGAGAGCTCAAGCTTTTCAGAGAGAAAGCAA 1845
|||
DB 567 CATGATATGACACCCAAAGTGGGAAACGAGAGCTCAAGCTTTTCAGAGAGAAAGCAA 508
|||
QY 1846 AGAGTAGCAATTCAGAGAGCCATTTTGAAGAGCCCCCAGTCATATCTATGAAAGCT 1905
|||
DB 507 AGAGTAGCAATTCAGAGAGCCATTTTGAAGAGCCCCCAGTCATATCTATGAAAGCT 448
|||
QY 1906 ACTTCATGCTTGAATTCGATTAAGAGACATTTCTTGGTCCATGAAGATGTGTC 1965
|||
DB 447 ACTTCATGCTTGAATTCGATTAAGAGACATTTCTTGGTCCATGAAGATGTGTC 388
|||
QY 1966 AAACACAGAACTTCATTTTATTCACACAGATTTGCAAGTGTGATGCAAGATGAA 2025
|||
DB 387 AAACACAGAACTTCATTTTATTCACACAGATTTGCAAGTGTGATGCAAGATGAA 328
|||
QY 2026 ATCATTTGCTTGAATTCAGAGTAAAGTACCCAGTGGTACCAATGATTTGCTTGT 2085
|||
DB 327 ATCATTTGCTTGAATTCAGAGTAAAGTACCCAGTGGTACCAATGATTTGCTTGT 268
|||
QY 2086 AACCTCATATATCTATTCGAAATGTGCATACAGAGACGCGTGTGCAAGACCAT 2145
|||
DB 267 AACCTCATATATCTATTCGAAATGTGCATACAGAGACGCGTGTGCAAGACCAT 208
|||
QY 2146 GATAACCCCAATGGGAAGCAAGAAATAATATCCAAAGAGAGAAAGAAAGAAA 2205
|||
DB 207 GATAACCCCAATGGGAAGCAAGAAATAATATCCAAAGAGAGAAAGAAAGAAA 148
|||
QY 2206 CTACAGAGAAATATGCAATAGTGTGAAAGCTGTGAAACTGTGCTAAGTCA 2265
|||
DB 147 CTACAGAGAAATATGCAATAGTGTGAAAGCTGTGAAACTGTGCTAAGTCA 88
|||
QY 2266 TAAGACATTTTCTTTTGTGTTTGTGACTACATATTTGCACTGAAGAGAAATGTTT 2325
|||
DB 87 TAAGACATTTTCTTTTGTGTTTGTGACTACATATTTGCACTGAAGAGAAATGTTT 28
|||
QY 2326 TATTAAAAAATCATACATT 2345
|||
DB 27 TATTAAAAAATCATACATT 8
|||

Search completed: February 15, 2003, 05:49:04
Job time : 3269 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 15, 2003, 04:58:20 ; Search time 88 Seconds

(without alignments)
10981.375 Million cell updates/sec

Title: AF133659

Perfect score: 4254
Sequence: 1 ATGGCCCTGCTCGCATGCA.....TATTAAAAAATCATACATT 2345

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xih
-O=acm2.1/USPTO.spool/AF133659/runat.10022003.155441.29523/app.query.fasta.1.2503
-DB=SPTRMBL.21 -QFMT=fastan -SUFFIX=epc -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=AF133659 @CGN 1.1 128 @runat.10022003.155441.29523 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp virus:*
16: sp bacteriopl:*
17: sp archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1852	43.5	606	5 Q9W0C5	Q9W0C5 drosophila

2	1727.5	40.6	728	10 Q91WML	Q91WML arabidopsis
3	1715.5	40.3	728	10 Q91F78	Q91F78 arabidopsis
4	1657.5	39.0	677	10 Q9M0G8	Q9M0G8 arabidopsis
5	1657.5	39.0	678	10 Q9FUT3	Q9FUT3 arabidopsis
6	1620.5	38.1	680	10 Q9M0G9	Q9M0G9 arabidopsis
7	1599	37.6	697	5 Q8T9W2	Q8T9W2 dicystoseli
8	1577.5	37.1	936	10 Q9S8N5	Q9S8N5 oryza sativ
9	1454.5	34.2	631	16 Q8ZD10	Q8ZD10 yersinia pe
10	1446.5	34.0	609	16 Q9ZDWO	Q9ZDWO rickettsia
11	1400	32.9	628	16 Q9YF88	Q9YF88 bruceella me
12	1399	32.9	592	16 Q8Y258	Q8Y258 ralsstonia s
13	1393.5	32.8	627	16 Q984W3	Q984W3 rhizobium l
14	1375	32.3	629	16 Q8UGH3	Q8UGH3 agrobacteri
15	1369	32.2	627	16 Q9ZRL2	Q9ZRL2 rhizobium m
16	1247	29.3	836	11 Q70595	Q70595 rattus norv
17	1241.5	29.2	842	11 Q9DC29	Q9DC29 mus musculu
18	1240.5	29.2	766	4 Q96ME8	Q96ME8 homo sapien
19	1239.5	29.1	896	4 Q9HMQ7	Q9HMQ7 homo sapien
20	1183	27.8	866	5 Q9VFE20	Q9VFE20 drosophila
21	1042	24.5	643	16 Q9A8N9	Q9A8N9 caulobacter
22	923.5	21.7	1025	5 Q25693	Q25693 plasmodium
23	915	21.5	704	5 Q9XUJ1	Q9XUJ1 caenorhabdi
24	908	21.3	947	5 Q08667	Q08667 plasmodium
25	799.5	18.8	320	5 Q95ST0	Q95ST0 caenorhabdi
26	725	17.0	642	17 Q9HSQ7	Q9HSQ7 halobacteri
27	709	16.7	581	16 Q9KDT6	Q9KDT6 bacillus ha
28	708	16.6	597	16 Q8YV35	Q8YV35 anabaena sp
29	706.5	16.6	593	16 P73239	P73239 synecocyst
30	700	16.5	832	16 Q9A2Y6	Q9A2Y6 caulobacter
31	699.5	16.4	578	16 Q99T13	Q99T13 stephylcococ
32	693.5	16.3	725	16 Q8XT23	Q8XT23 ralsstonia s
33	690.5	16.2	1243	16 Q9RDG5	Q9RDG5 streptomyce
34	687	16.1	585	16 Q31707	Q31707 bacillus su
35	682.5	16.0	735	2 Q9FV77	Q9FV77 rhizobium l
36	678.5	15.9	802	16 Q9RYG8	Q9RYG8 thermococcus
37	676.5	15.9	851	17 Q972N4	Q972N4 sulfobolus
38	675.5	15.9	594	2 P74940	P74940 thermotomaer
39	673.5	15.8	600	16 Q8U6S8	Q8U6S8 agrobacteri
40	673.5	15.8	863	17 Q97UJ9	Q97UJ9 sulfobolus
41	673	15.8	615	16 Q9CLD7	Q9CLD7 pasteurrella
42	672.5	15.8	546	16 Q8Y247	Q8Y247 anabaena sp
43	672.5	15.8	707	2 Q47258	Q47258 escherichia
44	672	15.8	708	2 Q93FG4	Q93FG4 manheimia
45	672	15.8	708	2 Q93313	Q93313 manheimia

ALIGNMENTS

RESULT 1

ID	Q9W0C5	PRELIMINARY;	PRT;	606 AA.
AC	Q9W0C5;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	CG7955 protein (GH20617p).			
GN	CG7955.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRATIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			

RA Bailew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bokora D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Butris K.C., Buam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kalp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclel J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svartsas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Mu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Niu C., Paclel J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; AE003472; AAF47525.1; -;
 DR EMBL; AY051556; AAK92980.1; -;
 DR FLYBase; FBgn0035244; CG7955.
 DR InterPro; IPR0013593; AAA_Arpase.
 DR InterPro; IPR001140; ABCtransporter.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transport; 1.
 DR SMART; SMO0382; AAA; I_TRANSPORTER; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 SQ SEQUENCE 606 AA; 66104 MW; 76C80500A1B62327 CRC64;

Alignment Scores:

Pred. No.: 3 31e-120 Length: 606
 Score: 1852.00 Matches: 360
 Percent Similarity: 76.92% Conservative: 110
 Best Local Similarity: 58.92% Mismatches: 131
 Query Match: 43.54% Indels: 10
 DB: 5 Gaps: 3

AF133659 (1-2345) x Q9W0C5 (1-606)

QY 358 ATCATAAAGCATGCTTTCTTATGTGTGGCCAAAGACGCGCATCTACAGACTAGA 417
 Db 1 MetLeuAArgAlaMetMeChalTyrlleTPrProLySGluAspProLeuValArgLysArg 20
 QY 418 GTTGCCATTTCCCTGGGATTTTGGGTGTGCAAAAGCCCATATATTTGGTTCCCTTC 477
 Db 21 ValGlyIleSerLeuGlyLeuLeuAlaGlySerLySLeuLeuThValCylValProPhe 40

QY 478 ATGTTAAATATGCTGTAGACAGCCTCAGCAGATGTCGGGAAACATGCTGAACCTGAGT 537
 Db 41 LeuHeuLySGlyAlaValAlaSpThMet-----ThrThrLeuAsnMetLeuAsp 55
 QY 538 GATGACCAAAATATAGTTTGGACACCATGGCAACAGACTTTCATGCTGTATGCTATCA 597
 Db 56 ThrIleProAlaValAlaLeuSerAlaAlaThrIleLeuMetLeuGlyTyGlyIleAla 75
 QY 598 AGACGTGGAGCTGCTTTTAAACGAATGCAATGCAATGCAATGCAATGCAATGCAATG 657
 Db 76 ArgAlaSerAlaAlaGlyPheAsnGluLeuArgAsnAlaValAlaPheAlaValAlaHis 95
 QY 658 AATTCATCCGAGAGATACCCAAATATGCTTCTTCATCTTCAACATCTGAGATCGGCT 717
 Db 96 HisSerIleArgLySleIleAlaSerAsnValPheLeuHisLeuHisAsnLeuAspLeuAla 115
 QY 718 TTTCACCTGAGACAGACAGCGGAGCTTTATCTAAGCTATTGACAGAGAACAGAGGCT 777
 Db 116 PheHisLeuAsnLySGlnThGlyAlaLeuSerLyThrIleAspArgGlySerArgGly 135
 QY 778 ATCAGTTTTCCTGAGAGCTTGTGATTTAATCTTCTCCATCATGTTGAAGTATG 837
 Db 136 IleAsnPheValLeuSerAlaMetValPheAsnIleValProThrIlePheGluLeuAla 155
 QY 838 CTTCGACGTGCTGTTTGTATTATCAAAATGCGGTCCAGATTGCTTGTGAACCTTGA 897
 Db 156 LeuValSerSerIleLeuGlyValIleCysGlyLeuAlaPheAlaGlyValSerMetGly 175
 QY 898 ACACCTGTACATACACAGCATTTACAGCTTGCAGCTCACCGGTGGAACATGATTAGA 957
 Db 176 CysValGlyIleTyAlaAlaTyThrLeuSerValThrGlnTrpArgThGlnPheArg 195
 QY 958 ATGAAATGAACAAAGAGATATGATGACAGTATGCTGCTATAGACTCAGCTGAT 1017
 Db 196 ValPheMetAsnGlnAlaGluAsnGluAlaGlyAsnLyAlaValAspSerLeuIleAsn 215
 QY 1018 TATGAACCTGGAAGTATTTAATATGAAGATATGAAGACACAGATATGATGATT 1077
 Db 216 TyrGlnTrpValLySlyPheAsnAsnGlnLySlyGlnAlaGlyCysTyAlaGlnVal 235
 QY 1078 TTGAAGACGTATGAGACTGCTTCACTTGAAGAAGTACCTTACTGCTGCTATGCTGA 1137
 Db 236 LeuLySlyTyGlnAlaAlaSerLeuLyThrSerSerLeuAlaLeuLeuAsnPhe 255
 QY 1138 GGTCAAGTCTATTTTTCAGGTGCTGTTTAAACAGCTTAATGAGTGCAGCAGAGGA 1197
 Db 256 GlyIleAsnAlaIlePheSerSerAlaLeuSerLeuIleMetValLeuAlaIleGly 275
 QY 1198 ATTGTGACAGGTACCTTACTGTTGAGATCTAATATGTAATGTAATGACTTTTTCAG 1257
 Db 276 IleAlaGlnGlyAsnMetThrValGlyAspLeuValMetValAsnAlaLeuLeuPheGln 295
 QY 1258 CTTTCATTAACCTTGAACCTTCTGGGAATCTGATATATAGAGACTAGACAAAGCTCATA 1317
 Db 296 LeuSerIleProLeuGlyPheLeuGlySerValTyArgGlnValArgGlnAlaLeuLeu 315
 QY 1318 GATATGAACACTGTTTACTCTCACTCAAGTATGACCAACCAATTAATGAACAAAGTATG 1377
 Db 316 AspPheArgAlaMetPheThrLeuMetAsnAlaAspSerSerIleGlnThrAlaAlaAsn 335
 QY 1378 GCATCTCCCTTGCATGACACACAGACAGACTACGCTGCGCTTGTATATGATGATATT 1437
 Db 336 AlaGlnProLeuPheValAspThrThrAsnSerSerIleGluPheArgAsnValSerPhe 355
 QY 1438 GAATACATTGAGGCCAGAAAGTCTTAGTGGAATATCTTTGAAGTCCCTGCGAGAAAG 1497
 Db 356 GluTyArgLupProGlyLySProllePheArgAspLeuSerPheThrIleProAlaGlyLys 375
 QY 1498 AAGTGGCATTTATAGAGAGTATGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 1557
 Db 376 AsnValAlaIleValIleGlyLySerGlySerGlySerGlySerSerMetValArgLeuLeuPhe 395
 QY 1558 CGCTTCTATGAGCTCAAAAGGATGACATTATCTGCTGATCAAAATATACAGATGTG 1617

```

Db 396 ArgPhePheGluProAnSerGlyValValLeuIleGlyGlnAspIleSerAlaVal 415
Qy 1618 AGCCGGAAGACCTTCGAGGCGGAGGAGTGTACCTCAGATGCTTCCTTCAT 1677
Db 416 AspleuGluSerLeuAlaGlyValIleAlaValAlaProGlnAspSerValLeuPheHis 435
Qy 1678 AAATCTATTATTACAACTCTTATATGAAACATCAAGTCTTCACCTGAGAAAGTAT 1737
Db 436 AsnThrIleGluIleAsnIleIleTyGlyAsnLeuSerIleSerHisAlaGlnValGln 455
Qy 1738 GCATGCGCAAAATTAAGCTTCATGATGATGATCTTCGATGATCCACATGATGATGAC 1797
Db 456 AsnAlaIleArgMetAlaAspLeuHisAspSerIleMetSerTrpProGlyGlnTyrSer 475
Qy 1798 ACCCAATAGGAGGAGGAGGAGCTCAAGCTTCAGAGAGAGAGAGAGAGAGAGATGAT 1857
Db 476 ThrGlnValGlyGlnValGlyLeuValLeuSerGlyGlyGlnValAlaIle 495
Qy 1858 GCAAGAGCCATTTGAAAGACCCCGCAGTCACTATGATGATGAGTACTTCATGCTTA 1917
Db 496 AlaArgAlaIleLeuIleAsnThrProIleLeuIlePheAspGlnAlaThrSerSerLeu 515
Qy 1918 GATTGATTAAGTGAAGACTATCTGTGTCCTGATGAGATGTCGTCAACACAGAACT 1977
Db 516 AspSerIleThrGlnHisAsnIleLeuGlnAlaLeuThrArgAlaThrSerGlyArgThr 535
Qy 1978 TCATTTTCATTCGACACAGATTTGTCAACAGTGTGATGATGATGATGATGATGCTTG 2037
Db 536 SerIleGlyIleAlaHisArgLeuSerThrValAlaAspAlaAspGlnIleLeuValLeu 555
Qy 2038 GATCAGGATTAAGTGAAGCGGAGCGTGTACCCAGCCATGTTGCTGTAAACCTCATAGT 2097
Db 556 GluAsnGlyArgValGlyGlnArgGlyThrHisSerGlnLeuLeu---ArgGlnAsnGly 574
Qy 2098 ATCATTTCAAGAAATGTGGCATACACAGACGCCGTGTGCAAGAACATGATACCCCAA 2157
Db 575 LeuTyrAlaArgLeuTyrGlnThrGlnThrGlnIlePhe-----AspProSer 590
Qy 2158 TGGGAGCAAGAAAGAAATATATATCCAAAGAG 2190
Db 591 ArgGlnIleAsnGlnGlnValAlaAlaValLeu 601

RESULT 2
Q9LVM1 PRELIMINARY; PRT; 728 AA.
AC Q9LVM1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ABC transporter-like protein (Half-molecule ABC transporter ATMs)
DE (Putative ABC transporter protein).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
RP SEQUENCE FROM N.A.
RC Sanchez-Fernandez R., Mari S., Dancis A., Rea P.A.;
RT "Functional half-molecule ABC transporters from Arabidopsis: the ATM
RT family."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.D., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene MCK7.14/AT5G58270 (GI:8777328).";
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL CC 1-SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AB019228; BAA96918.1; -.
DR EMBL; AF287699; AAG09829.1; -.
DR EMBL; AF360334; AAK26044.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_transporter_TM.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding, Transport.
SQ
SEQUENCE 728 AA; 80419 MW; A1PDDSC518E11B50 CRC64;

Alignment Scores:
Pred. No.: 1,44e-111 Length: 728
Score: 1727.50 Matches: 366
Percent Similarity: 65.91% Conservative: 125
Best Local Similarity: 49.13% Mismatches: 188
Query Match: 40.61% Indels: 66
DB: 10 Gaps: 10

AF133659 (1-2345) x Q9LVM1 (1-728)
Qy 39 CGCGGCGGCTGCTTCGAAAGCGCGGCACTC----- 71
Db 3 ArgGlySerArgPheValArgAlaProGlyLeuLeuLeuValLeuGlnPro 22
Qy 72 CGGATTCGATCCGCGCTTTAGTCTCTGTACCGGCTCAGGTCGCCAGTGGAGGCCACA 131
Db 23 GlnProIleIleProSerPheSerTyrSerLeuArg----- 34
Qy 132 TCACATCGGCGCTTGAGGAG-----CGTCCAGCTTACCATGAT 170
Db 35 SerAspTyrArgLeuHisAsnGlyPheSerAsnTyrIleArgArgAsnSerIleArgThr 54
Qy 171 TCCAGAGTCATTAAAGATAT----- 191
Db 55 SerProValIleAsnAlaPheLeuSerAspAsnSerProSerProSerProSer 74
Qy 192 -----CACATGCGCAGAGATTGGAGAAAGCAATTCAGAGACTTCTT 233
Db 75 ProIleArgPheValGlnArgSerSerMetLeuAsnGlyArgLeuPheSerThr----- 92
Qy 234 AGATGCTGCAAGAGCTCTCCAGGATATGCGCATGATGAGAAAGAGAGAGATGGCATGG 293
Db 93 -----SerThrProAsnProAspGlnThrThrIleThrIleSerIle----- 107
Qy 294 TCATGAGGAGGAGAGACTCCACAGACAGCCAAAGAGGATTAAAGATGTTGATCTCG 353
Db 108 -----LysThrThr-SerSerAspSerAspSerAlaMetAlaAspMet----- 121
Qy 354 GAAATCATTAAGACATGCTTTCTATGTGTGGCCCAAGACAGCCAGATTAAGAGC 413
Db 122 -LysIleLeuArgThrLeuAlaGlyTyrLeuTyrMetArgAspAsnProGlnPheArgPh 141
Qy 414 TTAGAGTTCGATTCGTCGGAGATTTTGGGTGGTGAAGAGCCATGATATATGTTGCC 473
Db 141 eArgValIleAlaAlaLeuGlyPheLeuValGlyAlaValValLeuValGlnValPhe 161
Qy 474 CTTGATGTTTAAATATGCTGTAGACAGCTCAACAGATGTGGGAAACATGCTGAACCT 533

```

```

Db 161 oPheLeuPheLysLeuAlaValAspTrpLeuAlaSerAlaThrGlyThrGlyAlaSerLe 181
Qy 534 GAGTATGACCAAAATATACATGTTGCACCATGGCAACA-----GCAGT 575
Db 181 uThrThrPheAlaAlaThrAsnProThrLeuThrValPheAlaThrProAlaAlaVala 201
Qy 576 TCTGATGGCTATGGTGTATGATCAAGAGCTGAGCTGCTTTTATTAAGAAATTCGAATGC 635
Db 201 lLeuLeIeIyGlyGlyIleAlaArgThrGlySerSerAlaPheAsnGluLeuArgThrAl 221
Qy 636 AGTATTTGGCAAGGTAGCCCAAGAAATTCATCCGAAGAAATAGCCAAAATGCTTTCTCCA 695
Db 221 aValPheSerLysValAlaLeuArgThrIleArgSerValSerArgLysValaPheSerHi 241
Qy 696 TCTTCACAACTGATCTGGGTTTTCACCTGAGCAGACAGACGGAGCTTTATCTAAGC 755
Db 241 sLeuHisAspLeuAspLeuArgLysIleLeuSerArgGluThrGlyGlyLeuAsnArgI 261
Qy 756 TATTGACAGAGAAACAGGGGTATCAGTTTGTCTGAGTGGCTGATTTAATCTCTCT 815
Db 261 eileAspArgGlySerAlaGlyAlaIleAsnPheIleLeuSerAlaMetValaPheAsnVala 281
Qy 816 TCCCATCATGTTGAAGATGATGCTGTGATGGTGTGTTTGAATTAACAATGCGGTGCCA 875
Db 281 lProThrIleLeuGluIleSerMetValSerGlyIleLeuAlaTyrlsPheGlyAlaAl 301
Qy 876 GTTTCCTTGGTAACTCTTGGAACCTTGGAACCTTGATACATACAGATTCACAGTTGCAGTAC 935
Db 301 aPheAlaTrpIleThrSerLeuSerValGlySerTyrlleValPheThrIleAlaValaTh 321
Qy 936 ACGGAGGAGAACTAGATTGAATGAATGAACAAAGACATATATGATGCAGTATGC 995
Db 321 rGlnTrpArgThrLysPheAsnGlyAlaIleAsnGlyAlaIleAsnGlyAlaIleAsnGlyAlaIle 341
Qy 996 TGTATAGACTCAGCTGCTGAATTTATGAACCTGTGAAGTATTTAAATGAAGAATATGA 1055
Db 341 gAlaIleAspSerLeuIleAsnTyrlGluThrValLysTyrlPheAsnAsnGluGlyTyrlG 361
Qy 1056 AGCAGACAGATATGATGATTTTGAACAGATGAGACTGCTTATGTAAGAAAGTACCTC 1115
Db 361 uAlaGluTyrlAspGlnPheLeuLysTyrlGluAspAlaAlaLeuGlnThrGlnAr 381
Qy 1116 TACTGTGGCTATGCTGATCTTGTCAAGTGTCTATTTTCACTGCGTTTAAACAGCTAT 1175
Db 381 gSerLeuAlaPheLeuAsnPheGlyGlnSerIleIlePheSerThrAlaLeuSerThrAl 401
Qy 1176 AATGTGCTCGCCAGCTCAGGAATTTGGCAGACCTTACTGTTGAGATCTAGTAAAT 1235
Db 401 aMetValLeuCyserGlnGlyIleMetAsnGlyGlnMetThrValaGlyAspLeuValMe 421
Qy 1236 GGTGAATGAGCTGCTTTTTCAGCTTTTCAATACCTCGTAACCTTCTGGGAACGTATATAG 1295
Db 421 tValAsnGlyLeuLeuPheGlnLeuSerLeuPheLeuAsnPheLeuGlySerValTyrlAr 441
Qy 1296 AGAGACTAGACAGACTCATATGATATGAACCTGTTTACTTACTCAAGAGTGAACAC 1355
Db 441 gGluThrIleGlnSerLeuValaAspMetLysSerMetPheGlnLeuLeuGluGluLysSe 461
Qy 1356 CCAAAATTAAGACAAAGTATGATCTCTCCCTTGATACACACACAGACAGTACCGT 1415
Db 461 rAspIleThrAsnThrSerAspAlaLysPheLeuValLeu-----LysGlyGlyAsnAl 479
Qy 1416 GGCCTTGATATATGCTATTTGAATACATTAAGGCGCAAGAAAGCTTATGGAATATC 1475
Db 479 egluPheGluAsnValaHisPheSerTyrlLeuProGluArgLysIleLeuAspGlyIleSe 499
Qy 1476 CTTGAAGTCCCTGAGAGAAAGATGGCCATTTAGAGAGTATGGGTACGGAAAG 1535
Db 499 rPheValaValProAlaGlyLysSerValaAlaIleValaGlyThrSerGlySerGlyLysSe 519
Qy 1536 CACATAGTGAAGCTATATATTCGCTTCATAGAGCTCAAAAGGTAGACATTAATCTTGC 1595

```

```

Db 519 rThrIleLeuArgMetLeuPheArgPheAspThrAspSerGlyAsnIleArgIleAs 539
Qy 1596 TGTCAAAATATATACAGATGAGCCTGGAAGCCTTCGAGGGCAGTGGAGTGTACC 1655
Db 539 pGlyGlnAspIleLysGluValaArgLeuAspSerLeuArgSerSerIleGlyValaIlePr 559
Qy 1656 TCAGGATGCTGTCTCTTCCATATATACATATTTATTAACCTCTTATATAGAAACATCAG 1715
Db 559 oGlnAspThrValLeuPheAsnAspThrIlePheHisAsnIleHisTyrlGlyArgLeuSe 579
Qy 1716 TGTTCACCTGAGGAAGTGTATGCAATGCGCAAAATTTAGCTGAGCTTATGATGCAATCT 1775
Db 579 rAlaThrGluGlnGluValaTyrlGluAlaAlaArgArgAlaAlaIleHisGluThrIleSe 599
Qy 1776 TCGAATGCCATGATATGATACCCCAAGTAGGGGAACGAGACTCAAGCTTTCAGAGG 1835
Db 599 rAsnPheProAspLysTyrlSerThrIleValaGlyGluArgGlyLeuLysLeuSerGlyG 619
Qy 1836 AGAAAGCAAAAGACTAGCAATTCGCAAGAGCCATTTGAAGAGCCCACTCATCTCTA 1895
Db 619 yGluLysGlnArgValaAlaLeuAlaIleArgThrPheLeuLysSerProAlaIleLeuLeuCy 639
Qy 1896 TGATGAAGCTACTTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1955
Db 639 sAspGluAlaThrSerAlaLeuAspSerThrThrGluAlaGluIleLeuAsnAlaLeuLy 659
Qy 1956 GGATGGGTCAACACAGAACTTATTTTCAATGTCACACAGATGTCACAGTGGTTGA 2015
Db 659 sAlaLeuAlaSerAsnArgThrSerIlePheIleAlaHisArgLeuThrThrAlaMetG 679
Qy 2016 TGCAGATGAATATCATTTGCTTGGATCAGGGTAAAGTACCGAACTGGTATCCACCATGG 2075
Db 679 nCyAspGlnIleValaValLeuGluAsnGlyLysValaValaGluGlnGlyProHisAspG 699
Qy 2076 TTGCTTGTCTAACCTCTATGATATCTTTCAGAAATGTGCTATCACAGACCGCTGT 2135
Db 699 uLeuLeuGlyLys---SerGlyArgTyrlaGlnLeuTrpThrGlnGlnAsnSerSerVa 718
Qy 2136 GCAGAACCATGAT 2148
Db 718 lAspMetLeuAsp 722

RESULT 3
Q9LF78 PRELIMINARY; PRT; 728 AA.
AC Q9LF78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Mitochondrial half-ABC transporter.
GN STAL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096387; PubMed=11158531;
RA Kishnir S., Babychuk E., Storozenko S., Davey M.W., Papenbrock J.,
RA De Rycke R., Engler G., Stephan U.W., Kispal G., Lill R.,
RA Van Montagu M.;
RT "A mutation of the Mitochondrial ABC transporter Stal leads to
RT dwarfism and chlorosis in the Arabidopsis mutant strick-";
RL Plant Cell 13:89-100(2001).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AJ272202; CAB97048.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransportTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.

```

DR SMART: SM0382; AAA, 1
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 SQ SEQUENCE 728 AA; 80394 MW; 29FESAAA2F77949C CRC64;

Alignment Scores:

Pred. No.:	9.81e-111	Length:	728
Score:	1715.50	Matches:	364
Percent Similarity:	65.64%	Conservative:	125
Best Local Similarity:	48.86%	Mismatches:	190
Query Match:	40.33%	Indels:	66
DB:	10	Gaps:	10

AF133659 (1-2345) x Q9LF78 (1-728)

```

QY CGCGCGCGCTCTTCGAAAAGCCGGCAGCTC----- 71
Db 3 ArgGlySerArgPheValArgAlaProGlyLeuLeuLeuCyArgValAsnLeuGlnPro 22
QY CGCGATTCTGATCCGGCCTTAGTCTGTGTTAGCGGCTCAGTCCGACAGTGAGGCCACA 131
Db 72 GlnProLysIleProSerPheSerTyrSerLeuArg----- 34
QY TCAACTCGCGCCCTTGGGAAC-----CGCTGAGCCTTACAGAT 170
Db 132 SerAspTyrArgLeuHisAsnGlyPheSerAsnTyrIleArgArgAsnSerIleArgThr 54
QY TCCAGAGTCATTAAAGATAT----- 191
Db 171 SerProValIleAsnAlaPheLeuSerAspAsnSerProSerProSerProSer 74
QY 192 -----CACATGCGACAGATTGGGAAAGCAATTTCAGACAGATTCTT 233
Db 75 ProIleArgPheValGlnArgSerSerMetLeuAsnGlyArgLeuPheSerThr----- 92
QY 234 AGATGCTGCAAAAGCTCTCCAGGTATGCGCCACTGATAGAAAAGACATTTGGCATGG 293
Db 93 -----SerThrProAsnProAspGlnThrThrThrIleThrLysGluIle----- 107
QY 294 TCATGACGAGAGGAGCTCCACACAGACCCCAAGAGGTTAAAGATGTTGATCTCG 353
Db 108 -----LysThrThr-SerSerAspSerAspSerAlaMetAlaAspMet----- 121
QY 354 GAAATCATATAAAGCAATGCTTTCTTATGTGTGCGCCAAAGACAGCCAGATTCACGAGC 413
Db 122 -LysIleLeuLeuArgThrLeuAlaGlyTyrLeuTyrMetArgAspAsnProGlnPheArgPhe 141
QY 414 TAGAGTTGCCATTTCCTCGGAGTTTGGGTGGTGCAGAGGCCATGAATATTGTGTTCC 473
Db 141 eArgValIleAlaAlaLeuGlyPheLeuValGlyAlaIleValLeuAsnValGlnValPr 161
QY 474 CTTGATCTTTAAATATATGCTAGACAGCTTCACACAGATGTCGGGAAACATGCTGAACCT 533
Db 161 oPheLeuPheLeuLeuAlaValAspTyrPheAlaSerAlaThrGlyAlaSerLe 181
QY 534 GAGTGATGCAACCAATACAGTTGCACACCATGCAACA-----GCAGT 575
Db 181 uThrThrPheAlaIleAlaThrAsnProThrLeuLeuThrValPheAlaThrProAlaAlaVala 201
QY 576 TCTGATTGGCTATGGTGTATGATCAAGAGCTGAGGCTTTTTTTAAAGAGTTGCAAAATGC 635
Db 201 lLeuIleGlyTyrGlyIleAlaArgThrGlySerSerAlaPheAsnGlnLeuAlaGlyThrAl 221
QY 636 AGTATTTGGAGAGTAGCCCAAGATTCAATCCGAAAGAAATGACCAAAATGCTTTTCCA 695
Db 221 aValPheSerIleValAlaLeuArgThrIleArgSerValSerArgIleValPheSerThrI 241
QY 696 TCTTCACAACTGGATCTGGGTTTTCACCTGACGACACAGACGAGCACTTATATTAAGGC 755
Db 241 sLeuHisAspPheAspPheLeuArgTyrHisIleLeuSerArgGlyuThrGlyLeuAsnAlaGlyI 261
QY 756 TATTGACAGAGCAAGGGGTATCAAGTTTGTCTGAGTGCTTTGGATTTAATCTTTCT 815

```

```

Db 261 eIleAspArgGlySerArgAlaIleAsnPheIleLeuSerAlaMetValPheAsnVala 281
QY 816 TCCCATCATGTTTGAATGATGATGCTTGTGAGTGGTGTGTTGTTATATCAAAATCGGTGCCA 875
Db 281 lProThrIleLeuGlnIleSerMetValSerGlyIleLeuAlaTyrLysPheGlyAlaI 301
QY 876 GTTTGCTTTGGTAAACCTTGGAACTGTGTACATACAGCATTCACAGTTGCAGTCAC 935
Db 301 aPheAlaThrIleThrSerLeuSerValGlySerTyrIleValPheThrLeuAlaValTh 321
QY 936 ACGGTGAGAACTAGATTTGAATAGAAATGAAACAAAGCATATATGATGACGTAATGC 995
Db 321 rGlnThrPArgThrLysPheSerIleValaMetAsnLysAlaAspAsnAspSerThrAr 341
QY 996 TGTATATGACTCACTGCTGAATTAATGAACCTGCAAGTATTTAATATGAATATGA 1055
Db 341 gAlaIleIleAspSerIleuLeuAsnTyrGlnThrValIleValTyrPheAsnAsnGlyArgI 361
QY 1056 AGCAGACAGATATGATGATTTTGAAGACGTATGACATGCTTCATTGAAAAGTACCTC 1115
Db 361 uAlaGlyLysTyrAspGlnPheLeuLysLysTyrGlyAspAlaAlaLeuGlnThrGlnAr 381
QY 1116 TACTTGGCTATGCTGAACCTTGTGCAAACTGCTATTTTACGTGCGTTTACAGCTAT 1175
Db 381 gSerLeuAlaPheLeuAsnPheGlyGlnSerIleIlePheSerThrAlaLeuSerThrAl 401
QY 1176 AATGATGCTCGCCAGTCAGGAAATGCGCAGGATCCCTTACCTTGTGAGATCATGTAAT 1235
Db 401 aMetValLeuCySerGlnIleIleMetAsnGlyGlnMetThrValGlyAspLeuValMe 421
QY 1236 GGTGAATGACCTGCTTTTTCAGCTTTTCATTACCCCTGAACTTTTGGGAATGTATATAG 1295
Db 421 tValAsnGlyLeuLeuPheGlnIleuSerIleuProLeuAsnPheLeuGlnYserValTyrAr 441
QY 1296 AGAGACTAGCAAGCACTCATATATATGAACACTTGTTTACTTACTCATCAAGTAGACAC 1355
Db 441 gGlnThrIleGlnIleSerIleuValAspMetLysSerMetPheGlnLeuLeuGlnGlyLys 461
QY 1356 CCAAAATTAAGCAAAAGTATGATGCTATCCCTTCAGATACACACACAGACAGTACCGT 1415
Db 461 rAspIleThrAsnThrSerAspAlaLysProLeuValLeu-----LysGlyIleValenI 479
QY 1416 GGCCTTGTATATATGCTATTTGAATATACATTTAGAGGCCAGAAAGCTTTAGTGAATATC 1475
Db 479 eGlnPheGlnAsnValHisPheSerTyrLeuProGlnArgLysIleLeuAspGlyIleSe 499
QY 1476 CTTTGAAGTCCCTGCGAGAAAGAAAGTGGCCATTGTAGAGGATGAGGTGACGGAAAG 1535
Db 499 rPheValIleProAlaGlyLysSerValAlaIleValGlyThrSerGlySerGlyLysSe 519
QY 1536 CACAATATGAGGCTATTTATTTGCTTATAGAGCTCAAAAGGTAAGCATTTATCTTGC 1595
Db 519 rThrIleLeuArgMetLeuPheArgPheAspThrAspSerGlyAsnIleArgIleAs 539
QY 1596 TGGTCAAAATATATCAAGATGTGAGCTGGAAGCCTTGGAGGGGACGTGGAGATGTGATCC 1655
Db 539 pGlyGlnAspIleLysGlyValaLargLeuAspSerLeuAspSerIleGlyValaValPr 559
QY 1656 TCAAGATGCTGCTCTTCCATATATCTATTTAACAACCTTTATATGAGAAATCAG 1715
Db 559 oGlnAspThrValLeuPheAsnAspThrIlePheHisAsnIleHisTyrGlyLysArgLeuSe 579
QY 1716 TGGTCACTGAGGAGATGTATGACATGCGAAATTAAGCTGAGCACTTCATGATGCAATTC 1775
Db 579 rAlaThrGlnGlnGlnValaTyrGlyAlaAlaArgAlaAlaIleHisGlnThrIleSe 599
QY 1776 TCGAATGCCACATGATATGACACCCCAAGTAGGGAAGAGAGATCAAGCTTTCAGAGAG 1835
Db 599 rAsnPheProAspLysTyrSerThrIleValGlyuArgGlyLeuLysLeuSerGlyGly 619
QY 1836 AGAAAAGCAAAAGATGCAATTCGAAAGAGCCATTTTGAAGAGACCCCACTACTCTTA 1895
Db 619 yGlyLysGlnArgValaIleAlaLeuAlaArgThrPheLeuLysSerProAlaIleLeuLeuCy 639

```

[illegible]


```

Qy 1210 ACCCTTACTGTTGAGAGATCTAGTAATGTAATGAGACTGCTTTTTCAGCTTTCATTATACC 1269
Db 368 GUMethrlnvalGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeuPro 387
Qy 1270 CTGAACCTTCTGGGAAGCTGTATATAGAGAGACTAGCAAGCACTCATAGATATGAAACCC 1329
Db 368 LeuTyPheLeuGlyGlyValTyArgGlnThrValGlnGlyLeuValAspMetLeuSer 407
Qy 1330 TTGTTACTCTCTACTAGAGTATAGACCCCAATTAAGCAAGGTAGTATGCACTCCCTT 1389
Db 408 LeuPheGlnLeuLeuGlnGlyLeuArgSerAspIleGlyAspLeuAspThrGlnThrLeu 427
Qy 1390 ---CAGATCACACACAGACAGACAGTACCGTGCCTTGTATATGTCATTTTGAATACAT 1446
Db 428 ProPheValLeuValArgGlyGlySerIleSerPheGlnValAsnValHisPheSerTyLeu 447
Qy 1447 GAGGCCCAAGAAAGTCTTATGTAATATCTTGAAGTCCCTGCGAGAAAGAAAGTGGCC 1506
Db 448 ProGlnArgGlyIleLeuAspGlyIleSerPheGlnValProAlaGlyLysSerValAla 467
Qy 1507 ATTGAGAGAGTAGTGGGTGAGGAAAGCAATAGTAGAGGCTATTTATTCGCTTCTAT 1566
Db 468 IleValGlySerSerGlySerGlyLysSerThrIleLeuArgMetIlePheArgPhePhe 487
Qy 1567 GAGCTCTCAAAAGGTAGCATTTATCTTGTGCTCAAAATATACAAGATGAGCTTGAA 1626
Db 488 AspThrAspSerGlyAsnValArgIleAspGlyGlnAspIleGlyGlnValThrLeuGln 507
Qy 1627 AGCCTTCGAGGGGAGTGGAGTGTAGCTCAGATGCTGTCTTCCATTAATCTAT 1686
Db 508 SerLeuArgSerGlyIleGlyValValProGlnAspThrValLeuPheAsnAspThrIle 527
Qy 1687 TATTACACCTTTTATATGAGAAACATGAGTGTGCTTCACTGAGAGAGATGATGCA 1746
Db 528 PheHisAsnIleLeuTyGlyAsnLeuSerAlaThrGlnGlnGlnValTyArgAlaAla 547
Qy 1747 AAATTAGCTGAGCTTCAATGATGCAATTTCTTCGATGCCACATGATATGACACCAAGTA 1806
Db 548 ArgArgAlaValIleHisAspThrIleMetCysPheProAspTySerThrAlaVal 567
Qy 1807 GGGGAACGAGACTCAAGCTTTTCAGAGAGAGAAAGCAAGTAGTCAATTCGAAGACC 1866
Db 568 GlyGlnArgGlyLeuMetLeuSerGlyGlyGlnArgValAlaLeuAlaArgAla 587
Qy 1867 ATTTTGAAGGACCCCACTTACTTATGATGTAAGGACTTCAATCGTATGATTCGATT 1926
Db 588 PheLeuTySerProAlaIleLeuLeuCysAspGlnAlaThrAsnAlaLeuAspSerLys 607
Qy 1927 ACTGAAGAGACTATCTTGTGTCATGAGAGATGTCACCAAGCAAGCACTTCTATTTTC 1986
Db 608 ThrGlnAlaGlnIleMetCysThrPheArgSerLeuAlaSerAsnArgThrCysIlePhe 627
Qy 1987 ATTGCACACAGATTGTCAACAGTGTGATGTCAGATGAATGATATTTGTCGATGAGGCT 2046
Db 628 IleAlaHisArgLeuThrThrAlaMetGlnCysAspGlnIleIleValMetGlnLysGly 647
Qy 2047 AAGGTAGCCGAACGTGTACCAACCATGCTTGTGCTTCAACCTCATATGATCTATTC 2106
Db 648 LysValValIleGlnLysGlyThrHisGlnValLeuLeu---GlnLysSerGlyArgTyAla 666
Qy 2107 GAAATGTGCATACACAGAGACGCGTGCAG 2139
Db 667 LysLeuThrThrGlnGlnHisSerThrLeuGln 677

```

```

GN AT428620.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mewes H.W., Lemcke K., Mayer K.P.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Sanchez-Fernandez R., Mari S., Dancis A., Rea P.A.;
RL "Functional half-molecule ABC transporters from Arabidopsis: the ATM
RT family."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AL161573; CAB81450.1; -.
DR EMBL; AF287698; AAC09828.1; -.
DR HSSP; P13569; INBD.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransportrTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transportr; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KM ATP-binding; Transport.
SQ SEQUENCE 680 AA; 76049 MW; 42169C4319FBE880 CRC64;

Alignment Scores:
Pred. No.: 3,73e-104 Length: 680
Score: 1620.50 Matches: 321
Percent Similarity: 71.32% Conservative: 129
Best Local Similarity: 50.87% Mismatches: 170
Query Match: 38.09% Indels: 11
DB: Gaps: 5

AF133659 (1-2345) x Q9M0G9 (1-680)
Qy 256 GTATGGCCATGATATGAAGAGACATGTTGCGATGCTCATGACAGAGACTTCAC 315
Db 57 IleThrProMetValAsnAlaArgValMetPhePheSerThrSerThrSerAlaProHis 76
Qy 316 ACAAGCCCAAAAGAGGTTAAAGATGTTGATCTCGGAAATCATMAAGCAATGCTT 375
Db 77 -----ProGlnLysIleAsnArgThrSerSerGlnAsnIleLeuArgMetIleSer 93
Qy 376 TCTTATGTTGGCCCAAGAGACAGGACAGATCTACAGCTGAGTGTGCAATTTCCGCTGGA 435
Db 94 SerTyLeuThrMetLysAspAsnProLysLeuCysPheArgValIleSerAlaPheAla 113
Qy 436 TTTTGGGTGTGCAAGGCCCATGAATATTGTGTTCTTCATGATGTTTAAATAGTCTGA 495
Db 114 CysLeuValGlyAlaLysPheLeuAsnValGlnValProPheLeuPheLysValAlaIle 133
Qy 496 GACAGCTTCAACCAAGATGTGGGAAACAATGCTGAACCTGAGTAGACCAACCAATACGTT 555
Db 134 AspThrPheSerSerSerSer-----PheValAspSerAsnProTyLeuVal 149
Qy 556 GCAACCATGGCAACA-----GCAGTTCTGATTTGGCTATGCTGATCAAGAGCTGAGCT 609
Db 150 AlaAlaPheAlaThrProSerSerValLeuIleGlyTyGlyIleAlaArgSerGlySer 169
Qy 610 GCTTTTAAAGAGTTCGAATATGAGATATTTGGCAAGGATGACCCGAATTCATTCGA 669
Db 170 SerAlaPheAsnGlnLeuArgThrSerValPheSerLysValAlaLeuArgThrIleArg 189

```

```

RESULT 6
Q9M0G9 PRELIMINARY; PRT; 680 AA.
AC Q9M0G9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ABC transporter-like protein (Half-molecule ABC transporter
DE ATW2).

```


QY 670 AGAATAGCCAAAATGCTTTTCTCATCTTCAACACCTGATCTGGGTTTTTCAACCTGAGC 729
 Db 190 ThrilleserArglyValleuSerArgleuHibAspleuAspLeuArgTyHileuAsn 209
 QY 730 AGACAGACGGGAGCTTTATCTTAAGCTTTTACAGAGGACAGAGGGGATATGATTTGTC 789
 Db 210 ArgAspThrGlyAlaLeuAsnArgIleIleAspArgGlySerArgAlaIleAsnThrIle 229
 QY 790 CTGAGTGGCTGATTTATCTTCTTCCATCATGTTGAAGTGAAGTGGTCTGATGGT 849
 Db 230 LeuSerAlaMetValPheAsnIleMetProThrIleLeuGluIleSerMetValSerCys 249
 QY 850 GTTTTGTATTAACAATGGGAGTCCAGTTGCTTTGTTGTAACCTTGGAAACCTTGATCA 909
 Db 250 IleleuAlaTyrlsPheGlyAlaValTyrlaLeuIleThrCysLeuSerValGlySer 269
 QY 910 TACACAGATTCACAGTTGTCACACAGGTGAGAACTAGATTAGATTAAGAAATGAC 969
 Db 270 TyrlleAlaPheThrIleuHlaMetThrGlnTrpArgIleTyrlaArgTyrlaMetAsn 289
 QY 970 AAACAGATTAATGATGAGGTAATGCTGTATGACTGATCTGAAATTTATGAACCTGTC 1029
 Db 290 GluAlaGluAsnAspAlaSerThrArgAlaIleAspSerLeuIleAsnTyrlGluThrVal 309
 QY 1030 AAGTATTTTATTAATGAAGAATATGAAGACACAGATATGATGATTTTGAAGACGTA 1089
 Db 310 LysTyrlPheAsnAsnGluAspTyrlGluAlaArgTyrlAspGlnLeuHlaSGluAsnTyrl 329
 QY 1090 GAGACTGCTTCAATGAAGAAGTACTACTGCTGCTATGCTGAACCTTGGTCAAGTCT 1149
 Db 330 GluAspAlaIleAlaLeuGlnSerArgLysSerPheAlaLeuAsnAsnGlyGlnSerPhe 349
 QY 1150 ATTTCAGTGCCTTTAACAAGCATATATGCTGCTCCGACAGTGGGAATTTGGAGCAGT 1209
 Db 350 IlePheSerThrAlaLeuSerThrAlaMetValLeuCysSerGlnGlyIleMetAsnGly 369
 QY 1210 ACCCTACTGTTGAGATCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
 Db 370 GlnMetThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeuPro 389
 QY 1270 CTGAACTTTCTGGGAAGCTGTATATAGAGAGATGAGACAGCACTCATAGATATGACAC 1329
 Db 390 LeuTyrlPheLeuGlyValIleValTyrlSerAspThrValGlnGlyLeuValAspMetLys 409
 QY 1330 TTGTTTACTCTACTCAAGGTGAGACCCCAATTAAAGCAAAAGTGAAGGCTCTCCCTT 1389
 Db 410 MetPheLysPheLeuGluGluArgSerAspIleGlyAspLysAspIleAspArgLysLeu 429
 QY 1390 ---CAGATCACACCAAGACAGACAGTACCGTGGCTTTGATATGATGATTTGATATGAT 1446
 Db 430 ProProLeuValLeuLysGlyGlySerIleSerPheGluAsnValHibPheSerTyrlLeu 449
 QY 1447 GAGGCGCAAGAAAGTCTTCTAGTGAATATCTTGAAGTCCCTGAGAGAAAGAGGCGC 1506
 Db 450 ProGluArgLysIleLeuAspGlyIleSerPheGluValProAlaGlyLysSerValAla 469
 QY 1507 ATTCTAGAGGTAGTGGTCTGAGGAGAAAGCATATGATGAGGCTATTTTCCTTCTAT 1566
 Db 470 IleValIleSerSerGlySerGlyLysSerThrIleLeuArgMetIlePheArgPhePhe 489
 QY 1567 GAGCGCTCAAAAGAGTACATTTATCTGCTGCTCAAAATATATACAAAGTGAAGCTGGAA 1626
 Db 490 AspValAspSerGlyAsnValLysIleAspGlyLysIleAspGlyValArgLeuGlu 509
 QY 1627 AGCTTGGAGGAGGAGTGGAGTGAATCTTCAAGATCTGTCTCTTCAATATATATAT 1686
 Db 510 SerLeuArgSerSerIleGlyValIleProGluAspThrValLeuPheAsnAspThrIle 529
 QY 1687 TATTACAACTCTTATATGAGAAATCAGTCTTCACTGAGAGAGTGTATGACGTGGCA 1746
 Db 530 PheIleAsnIleHibTyrlGlyAsnLeuSerAlaThrGluGluValTyrlAsnAlaAla 549

QY 1747 AATTAAGTGAAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
 Db 550 ArgArgAlaIleAlaIleHibAspThrIleMetLysPheProAspLysTyrlSerThrAlaVal 569
 QY 1807 GGGGAACGAGAGCTCAAGCTTTTCAAGAGAGGAAAGAAAGATGAGATTTGCAAGGCC 1866
 Db 570 GlyGluArgGlyLeuMetLeuSerGlyGlyGluLysGlnArgValAlaLeuAlaArgAla 589
 QY 1867 ATTTTGAAGACCCCGACATATCTATATGATGATGATGATGATGATGATGATGATGAT 1926
 Db 590 PheLeuLysSerProAlaIleLeuLeuCysAspGluAlaThrSerAlaLeuAspSerLys 609
 QY 1927 ACTGAAGACATATCTTGGTCCATGAAGATGCTGCTGAACAGACATCTTATTTTC 1986
 Db 610 ThrGluAlaGluIleMetLysThrLeuArgSerLeuAlaSerAsnArgThrCysIlePhe 629
 QY 1987 ATTGCACACAGATTTGTCAACAGTGTGATGATGATGATGATGATGATGATGATGAT 2046
 Db 630 IleAlaIleArgLeuThrThrAlaMetGlnCysAspGluIleLeuValMetGluLysGly 649
 QY 2047 AAGTAGCCGAGCGTGGTACCCACATGCTTGTCTTAACCTCATGATATCTTATCA 2106
 Db 650 LysValValGluLysGlyThrHibSGluValLeuLeuGlyLys--SerGlyArgTyrlAla 668
 QY 2107 GAAATGTGCATACACAGACAGCCGCTGTGCAG 2139
 Db 669 LysLeuTrpThrGlnGlnAsnSerLysLeuGlu 679
 RESULT 7
 Q8T9W2 PRELIMINARY; PRT; 697 AA.
 ID Q8T9W2;
 AC Q8T9W2;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE ABC transporter ABCB5.
 GN ABCB5.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxId=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Anjard C., Loomis W.F.;
 RT "Evolution of the ABC transporters of Dictyostelium."
 RT Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF466308; AAL74252.1; -
 SQ SEQUENCE 697 AA; 78531 MW; 304CCTBC2F344E2F CRC64;
 Alignment Scores:
 Pred. No.: 1,166-102 Length: 697
 Score: 1599.00 Matches: 312
 Percent Similarity: 72.02% Conservative: 123
 Best Local Similarity: 51.66% Mismatches: 153
 Query Match: 37.59% Indels: 16
 DB: 5 Gaps: 5
 AF133659 (1-2345) x Q8T9W2 (1-697)
 QY 316 ACAGACCCAAAGAGGTTAAAGATGTTGATCTCGGAAATCATTAAGCAATGCTT 375
 Db 108 ThrAspIleLysGlu-----ThrGlnAsnSerThrMetSerLeuLysThrValPhe 125
 QY 376 TCTTATGTGTGCGCCAAAGACAGCCAGATCTTCAGACTGAGTGTGCGATTTGCGTGGGA 435
 Db 126 TyrlTyrlLeuTrpProLysAspAsnAspSerLysIleLeuArgIleIleThrSerValLeu 145
 QY 436 TTTTGGGTGTCGCAAGGCCATGATATGTTGTTCCCTTCAAGTGTAAATATGCTGTA 495
 Db 146 LeuLeuLeuSerAlaLysValLeuThrValGlnIleProHibIlePheLysAspIleVal 165
 QY 496 GACAGCTCAACCAAGATGTCGGGAAACATGCTGAGATGATGACCAATAATACAGTT 555

Db	166	AspSerLeuThrTrpHrInGlnSerGluMetLeu-----	176
Qy	556	GCACCACTGGCAACAGCAAGTTCTGATGGCTATGGTATCAAGACSTGACSTGCTTTT	615
Db	177	---ThrLeuProLeuGlnGlyLeuLeuLeuAlaTyGlyAlaValAlaValSerGln	195
Qy	616	TTTAAAGCAAGTTGCAAAATGCAAGTATTTGGCAAGTGAACCCAGAAATTCATCCGAAGATA	675
Db	196	PheGlnGlnLeuAaArgGlnTrpHrLephSerLyValAlaIleAspAlaIleArgAspVal	215
Qy	676	GCCAAAATATGCTTTCTCCATCTTCCATCAACACCTGAACTGGGTTTTCACCTGACAGACAG	735
Db	216	SerCysSerThrPheIlyAspGluHisGlnLeuAspLeuThrPheHisIleuSerTrgIn	235
Qy	736	ACGGAGACTTATCTAAAGCTATTGACAGAGAAACAGGGATACAGTTTGTCTGTAGT	795
Db	236	ThrGlySerLeuSerArgIleIleLeaAspArgGlyValArgGlyIleAspHleuLeuAsn	255
Qy	796	GCTTTGGATTTAATCTTTCTTCCATCATGTGTTGAAGTATGCTTGTCAAGTGTGTTTG	855
Db	256	SerIleLeuPheHisValValProThrAlaPheGlnIleSerIleuValSerTrgValMet	275
Qy	856	TATTAACAATCCGGGCCCAATTTGGCTTTGGTGAACCTTGGAACCTTGCTGATACACA	915
Db	276	TyrThrThrLeuGlnGlyTrpGlnTyrSerAlaLeuSerIleuAlaThrIleAlaIleArgTrp	295
Qy	916	GCATTCAACAGTTGACAGTCAACGGGTGAGAACTAGATTGAATAGAAATGAACAACGA	975
Db	296	ValPheThrValIlyValThrIlyValTrpArgTrpGlnPheArgValIlyMetAsnIlyMet	315
Qy	976	GATATATATGCAAGTAAATGCTGCTATATAGTCTGCTGAATTTAATACTGTGAAGAT	1035
Db	316	AspAsnGlnIleAspAsnIlySmeMetAspSerIleuIleAsnHegIlnThrValIlyTrp	335
Qy	1036	TTTAATATGAAGAATATAGAACGACAGATATAGATGATTTTGAAGACGTATGACACT	1095
Db	336	PheAsnAspAspAlaLeuGlnValGlnArgTyrHisAsnTyrIleuIlySerGlnIlyAspLys	355
Qy	1096	GCTTCATTGAAGAAAGTACCTCTACTCTGGCTATGGTGAACCTTGGTCAAAAGCTGATTTTC	1155
Db	356	AlaSerIleuLysTrpHrThrSerSerLeuSerPheLeuAsnHegIlyGlnAlaIleuIlePhe	375
Qy	1156	AGTGTGGTTTAAAGCTATATATATGTGCTGCCAGTCAAGGAATTTGTGCAGATACCTT	1215
Db	376	SerLeuSerMetThrAlaMetMetIleMetAlaGlnIlyAlaGlnIlyAsnIleu	395
Qy	1216	ACTGTGTGAAGATCTAGTAAATGTGTAAAGACGTGTTTTCAGCTTTCATTAACCCGAC	1275
Db	396	SerValIlyAspLeuValLeuValAsnGlyIleuIleuPheGlnIleSerIleuProLeuAsn	415
Qy	1276	TTTCTGGAACTGTATATATAGAGACTAGACCTGATACATATAGATATGAACSTTGT	1335
Db	416	PheLeuGlnThrValIlyArgGlnIleIlyGlnSerIleuValAspMetAspHisIleuPhe	435
Qy	1336	ACTTACTCAAGTATAGACACCCAAATTTAAAGACAAAGTGAATGSCATCTCCCTTCAATC	1395
Db	436	SerLeuLeuAsnLeuAsnProIlyIleSerAspAsnIlyAspSerIlyProLeuIlyLeu	455
Qy	1396	ACACACAGACAGCAACCGTGGCTTTGATATATGTCATTTTGAATCAATTCAGAGGCCAG	1455
Db	456	-----GluAsnGlyThrIleValPheArgAspIleSerPheIlyTrpAsnAspSerVal	473
Qy	1456	AAAGCTCTTATGTGAATATCTTTGAAGTCCCTGACAGAAAGAAAGTGGCCATTTGAGA	1515
Db	474	GlnValIleuAsnAsnValSerPheGlnIlyGlnGlnIlyIlyArgIleAlaIleValGly	493
Qy	1516	GGTATGGGTGACAGGAAAAGACAAATAGTAGAGCTATATTTGCTTCTATAGACCTCAA	1575
Db	494	SerSerGlySerGlyLysSerThrLeuLeuArgIleuIleuTrpArgPheTrpAspValSer	513
Qy	1576	AAGGTATGACATTTATCTTGCTGCTGCAAAATATATACAGATGTAGAGCTGGAAAGCTTGG	1633
Db	514	SerGlySerIleGlnIleAspGlyGlnAspIleArgGlyIleGlnIleuIleuSerIleuTrp	533

Qy	1636	AGGCGAGTGGAGGTGGTACTCAGAGTGCCTGCTCTTCCATTAATACATTTATTATTCAC	1635
Db	534	LYSHSLIEGLIYALVALPROGLINAPRTHVALLEUPHEASAPRTHILETYTYASN	553
Qy	1696	CTCTTAATATGGAACATCAGTGGCTTCCAGTGAAGAAAGTATGACAGTGCAGAAATTAAGCT	1755
Db	554	ILEALATYGLYASNPFOANALATHTLYSGULGINVALGIUASALALALARGALALA	573
Qy	1756	GGACCTCATGTATGTCATTTCTTCAGATGCCACATGATATGACACCCAGTAGGGGACGA	1815
Db	574	HSLIEHSLIGLIVALILEULEAMENCLYSANGLYTYRSPRTHVALVALGIYLYARG	593
Qy	1816	GGATCTCAAGCTTTCCAGAGAGGAAAGAAAGAGTACAGCAATTCGAAGGCCATTTTGAAG	1875
Db	594	GLYLEUYSLEUSERGLYGLYGLYGLYSGINLARGVALSERILEALARGALALEUULYS	613
Qy	1876	GACCCCCAGTCATCTCTATGATGAAGCTACTTCATCGTTAGTTGCATTACTGAAGAG	1935
Db	614	ASPERPROILEVALPHETRYRASPGLUALATHINSESETLEUASPRTHCULYSGULYS	633
Qy	1936	ACTATTTCTGGTGCATGAAGATGTGTGTCAAACACAGACTTCATTTTTCATTGCACAC	1995
Db	634	LEULEMETRSPALALEUARGULTEUPHELYSGIYARGTRHTRHILEMELIALAHS	653
Qy	1996	AGATTTGCAACAGGCTTGATGCAATGAATCAATCATGTCTTGATCAGAGTAAG---GTA	2052
Db	654	ARGLEUSERHTRILEVALASPRALASPRGULILEVALLEGLYTHGLYLYLEILE	673
Qy	2053	GCCGACGTGTATCCACCACCATGTGGTTGCTTGCTTAACCTCATATGATATTCAGAAATG	2112
Db	674	LEUGLHARGGLYASNHISGLINGLINLEULEU---GLUTEUGLUGLYLYTYRARGSERMET	692
Qy	2113	TGGCATACACAG 2124	
Db	693	TRPLEUALAGLN 696	
RESULT 8			
Q9SNUS			
ID	Q9SNUS	PRELIMINARY;	PRT; 936 AA.
AC	Q9SNUS;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DR	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DI	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	ESNE A0067992 (C11433).		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatopsida; Magnoliophyta; Liliopsida; Poales; Poaceae;		
CC	Ehrhartoideae; Oryzaceae; Oryza.		
OX	NCBI_TaxID=4530;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=CV. NIPPONBARE;		
RA	Saaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC		
RT	clone: P0338C01."		
CC	Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.		
DR	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL, AP000391, BA88352.1, -.		
DR	InterPro, IPR003593, AAA_Alfase.		
DR	InterPro, IPR001140, ABCTransportrTM.		
DR	InterPro, IPR003439, ABC_transportr.		
DR	Pfam, PF00664; ABC_membrane; 1.		
DR	Pfam, PF00005; ABC_tran; 1.		
DR	ProDom, PD000006; ABC_transportr; 1.		
DR	SMART, SM00382; AAA; 1.		
DR	PROSITE, PS00211; ABC_TRANSPORTER; 1.		
DR	ATP-binding; Transport.		
QW	SEQUENCE 936 AA; 101685 MW; C52EE4AA5303ABEC CRC64;		
Alignment Scores:			
Score. No.:	3,73e-101	Length:	936
Freq. No.:	1577.50	Matches:	360

Alignment Scores:	
Pred. No.:	3.73e-101
Score:	1577.50
Length:	936
Matches:	360

Percent Similarity: 51.83%
Best Local Similarity: 38.79%
Query Match: 37.08%
DB: 10
Conservative: 121
Mismatches: 186
Indels: 261
Gaps: 12

AF133659 (1-2345) x Q95N05 (1-936)

QY 34 GCGGCGCGCGCGCTGTTTGGAAAAGCGCGACCTCCGCGATTCTGATCCGCTTTA 93
DB 27 ALaLaLaLaLaProLePhe-----ArgArgProProThrValProArgProLeu 43
QY 94 GTCTCT-----GTTACGGCTCAGTCCGACGTGGAGGCCACATCAACTCGGCGCTTG 147
DB 44 ProSerProLeuLeuGlyGlyPheGlyProAsn----- 54
QY 148 GGAACCGCTCGAGCTCAGACATTCCAGAGTCATTAAAGATACATGCGCAGAGATTG 207
DB 55 -----CysTrpValTyrPro 59
QY 208 GGAAGAGCAATTCAGACAGATTG-----TTAGATCTCGAAGGCTCTCCAGTATGG 261
DB 60 GlyAspGlyLysTyrAlaProPheGlyArgLeuSerCysPheMetSerArgSerThrTyr 79
QY 262 CCACGTGATGAAAAGAGACAGCATGTTGGCATGCTCATGCA----- 300
DB 80 ProProProArgAspVal-----ArgGlyHisAlaPheSerThrSerAlaAsnAla 97
QY 301 ---GAGAGAGACTCCACAGACCCAAAGAA-----GGGTTAAAGATGTT 345
DB 98 ValAlaValGlyLysSerSerAspAspLysValLysLysAspLysSerLysLysAspVal 117
QY 346 GATPACTCG-----AAATCATAAAGCAATGCTTCTTATGTTGTGCGCCAAA 393
DB 118 AspAspGlnIleAlaAspThrGlnIleLeuLysAsnLeuGlyLysTyrLeuLeuAsn 137
QY 394 GACAGGCGCAATTCAGACAGTACAGTTCGCTGGGATTTTGGGTGGCAAG 453
DB 138 AspSerProAspPheArgPheArgLeuIleLeuSerLeuGlyLeuValGlyAlaLys 157
QY 454 GCCATGATATTTGTTGCTCCCTTCAATTAATATGCTGAGACAGCTCCACAGATG 513
DB 158 ValIleAsnValGlnValProPheLeuPheLysLeuValAlaAspTrpLeuAlaLeu 177
QY 514 TCGGGAACATGCTGACCTGAGTATGACCAAAATACAGTTGCAACCATG----- 564
DB 178 AlagIylalagIuthrSerLeuAlaSerPheThrGluAlaAsnAlaThrLeuAlaLeu 197
QY 565 -----GCACACAGCATTTCTGATTGGCTGATGCTATCAGACAGCTGAGCTCTTTT 615
DB 198 PheAlaSerProAlaAlaValIleuIleGlyTyrGlyIleAlaArgSerGlyValSerAla 217
QY 616 TTTAACGAGTTCGAATTCAGATATTGGCAAGTACCCAGAAATTCAAATCCCAAGATA 675
DB 218 CysThrIleuLeuArgAsnAlaValPheSerLysValThrLeuArgAlaIleArgSerVal 237
QY 676 GCCAAA----- 681
DB 238 SerMetProSerArgGlyArgGlyThrArgArgArgProAspLeuArgGlyProAlaAla 257
QY 681 ----- 681
DB 258 AlaLeuAlaLysValAlaAlaAlaValAlaAspGluAlaGluValAlaGluPheValGlu 277
QY 681 ----- 681
DB 278 MetAlaValThrValLeuValGluArgArgTrpArgSerArgMetSerGlyGlyThrArg 297
QY 681 ----- 681
DB 298 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 317
QY 681 ----- 681

DB 318 SerArgAlaSerLysGlyAlaHisGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 337
QY 681 ----- 681
DB 338 GlyAlaGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 357
QY 681 ----- 681
DB 358 GluAlaAlaThrLeuValGlyAlaArgArgSerSerArgArgTrpAspGluPheValArg 377
QY 681 ----- 681
DB 378 GlnPheLeuAlaAspArgLeuValProAspLeuArgGlyIlePheLeuSerArgAspGln 397
QY 681 ----- 681
DB 398 ProIleProProThrProLysProAsnThrProLysSerGlyPheValProSerHisPro 417
QY 681 ----- 681
DB 418 IleProSerLeuGlnProAsnThrThrLeuLeuValLysLeuTrpValIleLeuProSer 437
QY 682 -----AATGCTTTTCATCTTCACAACTCGATCTGGGTTTCACTG 726
DB 438 GluLeuLeuLeuLeuGlnValPheSerHisLeuHisGluLeuAspLeuArgTyrHisLeu 457
QY 727 AGCAGACAGACGAGCTTATCTAAGGCTATGACAGACAGACAGACAGGCTATCAGTTT 786
DB 458 SerArgGlnThrIleAlaLeuAsnArgIleIleAspArgGlySerArgAlaIleAsnTyr 477
QY 787 GTCCTGAGCTTGGATTAATCTTCTCCATCATGTTGAAGTATGATGCTTGCAGT 846
DB 478 IleLeuThrValMetValPheAsnValAlaProThrIleLeuGlnIleGlyMetValSer 497
QY 847 GCGTTTGTGATTAACAATCCGCTGCCAGTTGCTTGGTAACTTGGAAACATTCGT 906
DB 498 SerIleLeuAlaTyrLysPheGlySerThrPheAlaTrpIleHisSerValSerValAla 517
QY 907 ACATACACAGCATTCACAGTGCAGTCACACGCTGAGAACTGATTTAGAAATGAATG 966
DB 518 ThrTyrIleAlaPheThrLeuAlaValThrGlnThrArgThrLysPheArgThrAlaMet 537
QY 967 AACAAAGCATATATGATGAGGTAATGCTGTAATGACTCACTGCTGAATTAAGAACT 1026
DB 538 AsnLysAlaAspAsnAlaSerSerThrValAlaValAspSerLeuLeuAsnTyrGluThr 557
QY 1027 GTGAGATTTTATATGAAGATATGAAGCACACAGATATGATGATTTTGAAGACG 1086
DB 558 ValLysTyrPheAsnAsnGlnGlnPheGluValGluLysTyrAspLysTyrLeuLys 577
QY 1087 TATGAGACTGCTTCAAGAAAGTACTGATCTGCTATGCTGAACCTTGTGCAAGT 1146
DB 578 TyrGluAspAlaAlaLeuLysThrGlnSerSerLeuAlaTyrLeuAsnPheGlyGlnAsn 597
QY 1147 GCTATTTTCAAGTCCGTTTAAACAGCTATATATGCTGCTGCCAGTACAGGAATTTGGCA 1206
DB 598 IleIlePheSerSerAlaLeuSerThrAlaMetValLeuSerSerLysGlyValMetSer 617
QY 1207 GGTACCTTACTGTTGGAAGTCTGTAATGCTGAATGACGCTTTTTCAGCTTTCAATTA 1266
DB 618 GlyAlaLeuThrValGlyAspLeuValMetValAsnGlyLeuLeuPheGlnLeuSerLeu 637
QY 1267 CCCGTAAGCTTCTGGGAAGTATATATGAGACAGTACAGACAGCTCATATGATGAAC 1326
DB 638 ProLeuAsnPheLeuGlySerValTyrArgGluSerArgGlnSerLeuIleAspMetLys 657
QY 1327 ACCCTGTTTACTTACTCAAGTATGACACCAATTAAGACAAAGATGATGATCTCC 1386
DB 658 SerMetPheGlnLeuLeuGlnGluLysProGlyIleLysAspGluProHisAlaGlnPro 677
QY 1387 CTTGAGATCACACACAGACAGCTACCGGCTTTGATATATGATGATTTGATATCATTT 1446
DB 678 LeuGlnPhe-----LysGlyGlyArgIleGluPheGluAsnValHisPheGlyTyrVal 695

Qy	1447	GAGGCGCAAGAAAGCTTACTAGTAATAATCCCTTGAAGTCCCGCAGGAAAGAAAGTGGCC	1506
Dy	696	Proglutinyb11leuLysGlyValatrhPheThrValProIadIyLysSerVala	715
Qy	1507	ATTGTAGAGGTATGTGGTCAGAGGAAAGACAAATGTAGAGCTATTATTTGGCTTCAT	1566
Dy	716	IleValGlyThrSerGlySerGlyLysSerThrIleuArgLeuLeuPheArgPhe	735
Qy	1567	GAGCTCAAAAGGCTAGCATTTATCTTGCTGTCAAAATATACAGATGTGACCTGGA	1626
Dy	736	AspSerSerSerGlySerIleArgIleAspGlyGlnAspIleArgGluValThrLeuAsp	755
Qy	1627	AGCCTTGAGAGGAGGAGGAGGAGTACCTCAGAGATGTGCTGCTCTCTCAATAATATCT	1666
Dy	756	SerLeuArgLysCyS11leGlyValValProGlnAspThrValLeuPheAsnAspThrIle	775
Qy	1687	TATTACAACTCTTATATGGAAACATCAGTGTCTTCACTGAGGAAGTATGACGTGCA	1746
Dy	776	LysHisAsnIleGlnTyrGlyArgLeuSerAlaThrAspGluGluValTyrThrAspAla	795
Qy	1747	AAATTAGCTGACCTTCATGATGCAATTTCTTGAATGCCACATGATATGACACCAAGTA	1806
Dy	796	ArgArgIalalalIleHisAspThrIleMetAsnPheProAspLysTyrAsnThrValVal	815
Qy	1807	GGGGAAGAGAGACTCAGCTTTCAGAGAGAGAAACCAAGAGTGAATTCAGACAGCC	1866
Dy	816	GlyGluArgGlyLeuLysLeuSerGlyGlyGluGlnThrArgAlaSerIleAlaArgVal	835
Qy	1867	ATTTTGGAAGACCCGCCAGTCATCTTATGATGAGAGCTACTTCATCGTTAATTCGAT	1926
Dy	836	PheLeuLysGluProSerIleLeuLeuCyAspGluAlaThrSerAlaLeuAspSerThr	855
Qy	1927	ACTGAAGAGACTATTTCTTGTCCTGACATGAGATGTGTCAAACACAGAACTTCTATTTTC	1986
Dy	856	ThrIuAlaSerIleLeuAsnSerLeuLysThrLeuSerValAspArgThrSerIlePhe	875
Qy	1987	ATTGCACACAGATTTGTCACAGTGTGATGTCAGATGAATCATTTGTCCTTATGAGGT	2046
Dy	876	IleAlaHisArgLeuThrThrAlaMetGlnCyAspGluIleIleValLeuGluAsnGly	895
Qy	2047	AAGGTAGCCGAACTGTGTATCCACCATGTTGCTTGAACCCCTCATAGTATCTATCA	2106
Dy	896	GluValValGluGlnGlyProHisAspPheLeuLeuSerLys---GlyGlyArgTyrAla	914
Qy	2107	GAATGTGGCATACACAGAGCAGC 2130	
Dy	915	GluLeuTyrSerGlnGlnAsn 922	

RESULT 9

08ZD10 PRELIMINARY; PRT; 631 AA.

AC 08ZD10;

RC 01-MAR-2002 (TReMBLrel. 20, Created)

RX MEDLINE=21470413; PubMed=11586360;

RA Patrick J.J., Weib B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,

RA Kendrick M.B., Seebhina M., James K.D., Churcher C., Mungall K.L.,

RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarrera A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Fellwell T., Hamlin N., Holroyd S., Uegels K., Kariyshev A.V.,

RA Leather S., Moulis S., Oyston P.C.F., Quail M., Rutherford K.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;

RT	"Genome sequence of Vesinia pestis, the causative agent of plague."
RL	Nature 413:523-527(2001).
DR	EMBL; A041452; CAC91388.1; -.
DR	InterPro; IPR001593; AAA_Atpase.
DR	InterPro; IPR001140; ABCtransprtTM.
DR	InterPro; IPR003439; ABC_transportr.
DR	Pfam; PF00664; ABC_membrane.1.
DR	Pfam; PF00005; ABC_tran.1.
DR	ProDom; PD00006; ABC_transporter; 1.
DR	SMART; SMO0382; AAA; 1.
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW	Hypothetical protein; Complete proteome.
SQ	SEQUENCE 631 AA; 71336 MW; 23JDA74D67B64F3B CnC64;
Alignment Scores:	
Pred. No.:	1.18e-92 length: 631
Score:	1454.50 Matches: 293
Percent Similarity:	67.73% Conservative: 110
Best Local Similarity:	49.24% Mismatches: 179
Query Match:	34.19% Indels: 13
DB:	16 Gaps: 3
AF133659 (1-2345) x Q8ZDI0 (1-631)	
QY	340 GATGTGATCTACGGAAAAATCATAAAAGCAATCCTTTTATATGTGTGGCCCAAAGACAGC 399
DB	35 AsnMetAspArgSerArgIleLeuLysPheLeuLeuProTyrLeuTrpProLysAspAsn 54
QY	400 CCAGATCTCACGAGCTAAGATGGCCATTTCGCGCGGATTTTTGGTGATGCAAAAGGCATG 459
DB	55 ProLysLeuAlaGlyTyrTyrLeuIleIleIleAlaLeuPhePheMetValIleAlaLysIleSer 74
QY	460 AATATGTGGTTCCTTCATGTTTAATAATGCTGTAGACAGCCTCAACAGATGTGGGA 519
DB	75 ThrThrLeuValProLeuAlaTyrLysSerMetValAspThrLeuSerSerCjLuSnaLa 94
QY	520 AACATGCTGAAACCTGAGTAGTGCACCAAATACAGTTGCCAACATGGCAACAGCACTTGTG 579
DB	95 LysMetLeu-----AlaIleProIleSerLeuLe 104
QY	580 ATTGCGATGGGTGATCATCAAGAGCTGGAGCGTCTTTTAAACGAAGTTCGAAATGCAGTA 639
DB	105 IleAlaTyrGlyValAlaIlaArgValIleSerIlaSerLeuPheGluGluLeuAlaRhsValMet 124
QY	640 TTGTGCAAGTGGCCCGCAATTCATCCGAGAAGATGCCMAAAATGTCTTCTCCACTT 699
DB	125 PheValHisValSerGlnSnaAlaIleAlaTrpArgLeuLeuGlyLeuAlaArgValPheArgGlnLeu 144
QY	700 CACAACCTGATCTGGGTTTTCACCTGAGCAGACAGACGGAGCTTTATCTAAGCTATT 759
DB	145 HisAlaLeuSerLeuAlaArgPheHisLeuGluArgGlnThrGlyLeuSerLeuSerIle 164
QY	760 GACAGAGCAACAAAGGGTATCACGTTTTCCTGAGCGCTTGTGGTATTAATCTCTCC 819
DB	165 GluArgGlyThrGlnAlaValIleSerThrValLeuSerArgLeuLeuPheSerIleLeuPro 184
QY	820 ATCATCTTTGAAGATGATCTGTCACTGGGTGTTTGTATCAAAATGGCGTGCCACGTT 879
DB	185 IleLeuPheGluIleThrLeuValSerValIleMetTrpArgLeuLeuSerGlyTyrPhe 204
QY	880 GCTTTGGTAAACCTTGGACACCTGGTATCATACACAGATTCACAGTTGCAGTCAACGG 939
DB	205 AlaLeuAlaIleLeuValThrValSerCyStyIleLeuPheTh ^g ValMetAlaValGly 224
QY	940 TGGAGACTAGATTATGAATAAATGAACAAAGCAGATAAATGATGTCAGATTAATCTCT 999
DB	225 TrpArgThrArgPheArgGlnLeuAsnLysAlaAsnAlaAspAlaAsnThrLysSer 244
QY	1000 ATAGATCTACGCTCGATTTATGAACCTGTGAGAGTATTTTAAATATGAAGAATGATACACA 1059
DB	245 IleAspSerLeuLeuAsnLysArgGlnThrValLysTyrPheGlyAsnGluAspPheGlnAla 264
QY	1060 CAGAGATATGATGGATTTTGAAGACGTTAGACACTGCTTCAATTGAAGAATGACTTACT 1119

Dd	265	284
		GUATGpHeaSnLeuSerArgInLeuTyGtIuYrAlaAlaLeuLysAsnInIleSer										
Qy	1120	CTGGGTAATCGGAACCTTTGGGCAAAAGCGTAATTTGAAGTGCAGTTTAACAGCTAATATG	1179	
Dd	265	PheThrValLeuSerPheGlyGlnThrAlaIleIleSerValGlyLeuIleValMetMet	304	
Qy	1180	GTGCTCGCCACTCGAGGAATTTGGCCAGGATACCCTTACTGTTGGAGATCTAGTAATGGTG	1239	
Dd	305	AlaMetAlaIaInGInGlyIleValGInGInGlyMetThrIleGlyYAspPheValLeuVal	324	
Qy	1240	AATGACGCTGCTTTTACGCTTTTCAATTAACCCCTGAACCTTCTGGGAACGTATATAGAAG	1299	
Dd	325	AsnAlaTyrLeuLeuGlnLeuTyGlnProLeuAsnPheGlyPheIleTySerGlu	344	
Qy	1300	ACTAGACAGACACTAGATAGATGAACACCTTTTAACTCTACTCAAGGTATACACACCA	1359	
Dd	345	IleArgGlnAlaLeuIleAspMetGluAsnMetLeuAspLeuMetValLysGlnGlu	364	
Qy	1360	ATTAAAGACAAAGTATGAGCATCTCCCTTCAGATCACACACAGACACTACCCGTGCC	1419	
Dd	365	IleThrAspArgProAspAlaLeuProLeuGlnLeuThr-----LysGlyValuValArg	382	
Qy	1420	TTTGATATATGTGCTTTTGAAATACATTAGAGGCCCAAAAGTCCTTATAGTGAATATCTTT	1479	
Dd	383	PheAspAlaValSerPheSerTyAspProIArgArgProIleLeuAsnAsnValSerPhe	402	
Qy	1480	GAAGCGCCCTGCAGGAAGAAAGGCGCATTTGTGAGGATAGTGGCGGGGAAAGACGA	1539	
Dd	403	ThrIleProGlyLysThrValAlaIleValGlyAlaSerGlyValAlaLysSerThr	422	
Qy	1540	ATAATGAGGCTATATTATTCGCTCTATAGACCTCAAAAGGGTAGACATTATCTGCTGCT	1599	
Dd	423	LeuAlaArgLeuLeuPheArgPheTyAspValThrAlaGlyAlaValTyLeuAsnAsp	442	
Qy	1600	CAAAATATACAGATGTGTGACCTTGGAAAGCTTTCGAGGGCGAGTGGAGTGTACTTAG	1659	
Dd	443	GlnAspIleArgGlyValIleThrGlnSerSerLeuArgGluAlaIleGlyLysValProGln	462	
Qy	1660	GATCGTGCCTCCCTCCATTAATACATTTATTAACACCTCTTATATAGGAAACATCACTGCT	1719	
Dd	463	AspThrValLeuPheAsnAspThrLeuArgTyAsnIleGlyTyGlyLysThrAspSer	482	
Qy	1720	TCACCTGAGGAAGTGTGTGAGTGGCCAAATTTGCTGCACTTCATGATGCAATCTTTCGA	1779	
Dd	483	ThrAspGlnGlnIleGluAlaAlaIleLysLeuAlaIleIleIleGlnPheIleIleSer	502	
Qy	1780	ATGCCACATGATATGACACCCCAAGTAGGGAAACGAGACCTCAACCTTTCAGAGAGAA	1839	
Dd	503	LeuProAspArgTyGlyGlnThrArgValGlyGlnArgGlyLeuTyLeuSerGlyLysIleu	522	
Qy	1840	AAGGAAGAGTACGAAATTTGGAAGGCCATTTTGAAGACCCCCCACTACTCTATATAT	1899	
Dd	523	LysGlnArgValAlaIleAlaIleArgThrIleLeuLysLysProSerIleLeuValPheAsp	542	
Qy	1900	GAAGCTACTTCATCGTATGATTCGATTACTGAAAGACATTCCTTGTCGCTCAAGAGAT	1959	
Dd	543	GlnAlaIleThrSerAlaLeuAspThrIleIleThrGlnArgGlnIleGlnSerIleLeuArgGlu	562	
Qy	1960	GTGGTCAAAACACAGACTTTATATTTTCATTTGACACACAGATTGTCAACAGTGGTATGCA	2019	
Dd	563	ValSerArgAspThrIleThrIleLeuValIleAlaIleIleArgLeuSerThrIleIleLeuAsp	582	
Qy	2020	GATGAATCATCTGCTTGGATCAAGGGTAAAGTAAAGCCGAACGTGGTAAACCAACATGGTTTG	2079	
Dd	583	AspIleIleIleValLeuGlnAlaGlySerIleValGlnArgGlyArgIleAspLeuIleu	602	
Qy	2080	CTTGCTAACCCCTCATAGTATCTATTCATGAGAAATGTGCGCATACAG	2124	

ID	Q9ZDM0	PRELIMINARY;	PRT;	609 AA.
AC	Q9ZDM0;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Mitochondrial transporter ATP1 precursor (ATM1).			
GN	RP205.			
OS	Rickettsia prowazekii.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;			
OC	Rickettsiaceae; Rickettsiaseae; Rickettsia.			
OX	NCBI_TaxID=782;			
XP	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=MADRID E;			
RX	MEDLINE=99039499; PubMed=9823893;			
RA	Anderson S.G.E., Zomorodpour A., Andersson J.O.,			
RA	Scheritz-Porten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,			
RT	Ericksen A.-S., Winkler H.H., Kurland C.G.;			
RT	"The genome sequence of Rickettsia prowazekii and the origin of			
RT	mitochondria."			
RL	Nature 396:133-140(1998).			
CC	-I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY			
CC	(ABC TRANSPORTERS).			
DR	EMBL; AJ235270; CAAL4670.1; -.			
DR	HSSP; P13369; INBD.			
DR	InterPro; IPR003593; AAA ATPase.			
DR	InterPro; IPR001140; ABCtransportM.			
DR	InterPro; IPR003439; ABC transportr.			
DR	Pfam; PF00664; ABC_membrane; 1.			
DR	Pfam; PF00005; ABC_tran; 1.			
DR	ProDom; PD00006; ABC_transport; 1.			
DR	SMART; SM00382; AAA; 1.			
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.			
KW	ATP-binding; Transport; Complete proteome.			
SEQ	SEQUENCE 609 AA; 68668 MW; 1126067A0FED8243 CRC64;			
Alignment Scores:				
	Pred. No.:	4.21e-92	Length:	609
	Score:	1446.50	Matches:	294
	Percent Similarity:	68.67%	Conservative:	118
	Best Local Similarity:	49.00%	Mismatches:	170
	Query Match:	34.00%	Indels:	18
	DB:	16	Gaps:	6
AF133659 (1-2345) x Q9ZDM0 (1-609)				
QY	370 ATGCTTCTTAATGTGTGGCCAAAGACGAGCCAGATCTAGAGCTAGAGTTGCCATTGCG	4228		
DB	23 LeuLeuThrTyrLeuTyrSerTyrAspPhe--AspIleArgLeuArgIleValThrSer	41		
QY	430 CTGGATTTTGGGTGGTGGCAAGGCAATGATATGTGGTTCCTTCATGTTAAATAT	489		
DB	42 LeuLeuCysLeuValIleAlaIleValIleLeuIlePheValProlIleValTyrLeuTyr	61		
QY	490 GCTGTGACAGCCTTCAACCAAGATGCGGAAACATGCTGAACCTGATGTGACCAAAAT	5498		
DB	62 IleIleAspGlyLeuAsnGln-----AsnLeuSer-----	71		
QY	550 ACAGTGCAACCATGGCAACGACGATTCTGATTGGCTAGTGGTATCAAGAGCTGAGCT	609		
DB	72 -----LeuSerValLeuIleGlyValIleIleGlyTyrGlyGlyThrIleLeuAla	89		
QY	610 GCTTTTTCAGAAAGTTCCAAATGACAGTATTTGGCAAGTAGCCCAAGATTCAATCCGA	6658		
DB	90 GlnIlePheSerGluLeuArgAsnIleIlePheSerIleValGlyCysGlnAlaThrArg	1090		
QY	670 AGAATGCAAAATGCTCTTCTTCATCTTCACACACCTGATGCTGGATTTTCACCTGAGC	7299		
DB	110 LeuValAlaLeuAsnValPheIleHisIleMetHisAsnLeuSerMetArgPheHisIleThr	1299		
QY	730 AGACAGACGGAGGCTTATCTTAAGGCTATTGACAGAGAAACAAGGGTATCAGTTTGTG	7899		
DB	130 ArgValThrGlyValLeuSerArgSerIleGluArgGlyThrIleValGlyIleGluAlaVal	1499		

Db 33 LeuArgMetArgValValTrrPalatThrPheTyrLeuValLeuSerIleValLeuIle 52
 QY 466 GTGGTCCCTTCATGTTAAATATGCTGTAGACAGCTCAACAGATGCGGAAACATG 525
 Db 53 LeuValProTyrPhePheLeuTrrValThrAsnAlaLeuAsn-----GlyGlnLeu 69
 QY 526 CTGAACCTGAGTANTGCAACAAATACAGTTGCACACAG-----GCAACAGCACTT 576
 Db 70 -----HisAlaProTyrTrrIleProValValLeuValGlyAlaValMetLeu 85
 QY 577 CTGATGGCTATGGTGTATCAAGAGCTGAGAGCTCTTTTATCAAGATCGAATGCA 636
 Db 86 ValLeuAlaTyrAsnAlaIleValIleValGlnIleValLeuAsnGlnLeuArgAspAla 105
 QY 637 GTATTGGCAAGTAGCCCAAGATTCATCCGAAGATAGCCAAATAATGCTCTTCAT 696
 Db 106 LeuPheAlaSerValGlyGlnTrrValAlaValArgIleLeuAlaTrrTyrLeuPheValHis 125
 QY 697 CTTCACAACCTGGATCTGGGTTTTCACCTGAGCAGACAGAGGAGCTTATCTTAAGCT 756
 Db 126 MetHisGlnLeuSerLeuValArgPheHisLeuGlnIleArgTrrGlyLeuSerArgVal 145
 QY 757 ATTAGACAGAGCAACAGGGGTATCATGTTTGCCTGAGTGTCTTGTATTAATTAATCTTCT 816
 Db 146 IleGlnIleArgGlyThrIleArgIleGlnIleValIleArgPheThrIleLeuAsnThrLeu 165
 QY 817 CCATCATGTTTGAAGTATGCTGTGAGTGTGTTTGTATTAACAATGCGGTGCCAG 876
 Db 166 ProThrIleLeuGlnIlePheAlaLeuThrAlaValIlePheAlaPheAlaTrrGlyLeuSer 185
 QY 877 TTTCCTTGGTACCCCTTGGAACCTTGATACATACAGCATTCACAGTTGCAGTCA 936
 Db 186 TyrLeuValValValAlaAlaIleThrValTrrPheTrrIleArgAlaSer 205
 QY 937 CGGAGGAACTAGATTGAATAGAAATGAAACAAGATATGATGACAGTAATGCT 996
 Db 206 AspTrrArgIleAsnIleHisArgGlnMetCysAspSerAspThrAspAlaAsnThrLys 225
 QY 997 GCTATAGACTCATCTGTAATTAATGAACCTGTGAAGTATTTAATAAGAAAGATATGAA 1056
 Db 226 AlaIleAspSerLeuLeuAsnPheGlnThrValLysTrrPheGlyAsnGlnAlaMetGln 245
 QY 1057 GCACAGAGATATGATGATTTTGAAGACGATAGAGCTGCTTATTCGAAAAGTACCTCT 1116
 Db 246 AlaLysArgPheAspGlyAlaMetAlaArgTrrGlnLysAlaAlaIleThrGlnThrTrrThr 265
 QY 1117 ACTGTGGCTATGCTGAACCTTGTGCAAGGCTATTTTTCAGTGTGGTTTAACGCTATA 1176
 Db 266 SerLeuGlnTrrPheAsnPheGlnIleValIlePheGlyAlaGlyMetAlaIleVal 285
 QY 1177 ATGTGTCTGCGCACTGAGGAAATGTGCGAGGTACCTCTTACTGTGGAGATCTAGTATG 1236
 Db 286 MetValMetSerAlaMetGlnValGlnLysGlyThrGlnSerLeuGlyAspPheValPhe 305
 QY 1237 GTGAATGACTGCTTTTCAAGCTTTCATTAACCTGTAACCTTTCGGAACGTATATAGA 1296
 Db 306 IleAsnAlaLeuLeuMetGlnLeuSerIleProLeuAsnPheIleGlyPheIleTrrArg 325
 QY 1297 GAGACTAGACAGCATCATAGATATGAACACCTGTTTACTTACTCAAGATAGACACC 1356
 Db 326 GlnIleArgGlnGlnLeuThrAspIleGlnGlnMetPheAspLeuAspValLysGln 345
 QY 1357 CAATTAAAGACAAGATGATGCTCCCTTCAGATCAACACAGACAGTACGCTAG 1416
 Db 346 GlnValLysAspLysProGlyAlaProAlaLeuLysVal-----AspSerGlyAlaIle 363
 QY 1417 GCCTTTGATATGTCATTTTGAATCATTTGAGGCGCAAGAAAGTCTTACTGGAATATCC 1476
 Db 364 SerPheLysAspValHisPheAlaTrrAspProGlnIleArgProIleLeuArgGlyIleSer 383
 QY 1477 TTGGAAGTCCCTGAGGAAAGAAAGTGGCATTTAGAGAGTGTGGTTCAGGGAAGC 1536
 Db 384 PheAspValProAlaGlyLysThrValAlaIleValGlyProSerGlyAlaGlyLysSer 403

QY 1537 ACAATAGTAGGCTATTATTTGCTTCTATAGCCTCAAAAAGGTAGCATTTATCTTGTCT 1596
 Db 404 ThrIleSerArgLeuLeuPheArgPheTrrAspIleGlnSerGlySerIleThrIleAsp 423
 QY 1597 GGTCAAAATATACAGATGTGAGCCTGGAAGCCTTGGAGGGGAGGAGTGGTACT 1656
 Db 424 GlyGlnAspValArgAspValTrrGlnIleSerLeuArgLysPheIleGlyMetValPro 443
 QY 1657 CAGGATGCTGCTCTTCATTAATACTATTATTAACAACCTTATTAAGAAACATCAGT 1716
 Db 444 GlnAspThrAlaLeuPheAsnAspThrIleAlaTrrAsnIleArgTrrGlyArgThrAsp 463
 QY 1717 GCTTTCACCTGAGAGTGTATGACAGTGGCAAAATTAAGTCTGCACTTCATGATTCCT 1776
 Db 464 AlaSerGlnIleAspValGlyLysAlaAlaGlnLeuAlaGlnIleAlaGlyPheIleLys 483
 QY 1777 CGAATGCCATGATGTATGACACCCAGTGGGGAAGAGAGACTTCACAGTTCAGAGGA 1836
 Db 484 HisLeuProAspGlyTrrLysSerMetValGlyGlnArgGlyLeuLysSerGlyGly 503
 QY 1837 GAAAAGCAAGAGTAGCAATTTGCAAGAGCCATTTTGAAGAGCCCGCATCATCTCTAT 1896
 Db 504 GlnLysGlnArgValAlaIleAlaIleAlaArgThrIleLeuLysAlaProProIleLeuIle 523
 QY 1897 GATGAAGCTACTTCATGCTTATGATTCGATTACTGAAGACATATTCTTGGTCCATGAG 1956
 Db 524 AspGlnAlaThrSerAlaLeuAspThrAlaThrGlnIleGlnIleGlnSerAlaLeuAsp 543
 QY 1957 GATGTGTCAAACAGCAAGATCTTCAATTTTCATTCACACAGATTTGCAACAGTGTGAT 2016
 Db 544 IleValSerArgGlyArgThrThrLeuValIleAlaHisArgLeuSerThrValIleGly 563
 QY 2017 GCAGATGAATCATTTGCTTGATGATGAGGTAGGTAGCGAGTGTACCCACATGCT 2076
 Db 564 AlaAspGlnIleIleValLeuLysAspGlyLeuIleAlaGlnArgGlyThrHisArgHis 583
 QY 2077 TTGCTTGCTAACCTCATATGATTCATTCAGAAATGTGGCATACACAGAGCCGTGTG 2136
 Db 584 LeuLeu---AspGlnLysGlyLeuTrrAlaSerMetTrrAspArgGln----- 598
 QY 2137 CAGAACCATATATACCCCAATAGGGAAGCAAGAAAGAAATATTCACAAAGAGAGGA 2196
 Db 599 -----ArgGlnAlaSerGlnAlaGlnIleArg 607
 QY 2197 AGAAGAACTACAGAA 2214
 Db 608 LeuArgGlnValArgGln 613
 RESULT 12,
 08Y258
 AC 08Y258; PRELIMINARY; PRT: 592 AA.
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Probable composite ATP-binding transmembrane ABC transporter protein.
 GN RSC0478 OR RSC04417.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 Arlat M., Billault A., Broctier P., Camus J.C., Catolico L.,
 Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 Gaspin C., Layte M., Moisan A., Robert C., Saurin W., Schlex T.,
 Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 Weisenbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"
 RL Nature 415:497-502(2002).
 DR EMBL: AL646059; CAD14006.1; --
 DR InterPro: IPR003593; AAA_Artpase.
 DR InterPro: IPR001140; ABCtransportTM.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transportr; 1.
 DR SMART: SM00382; ABC_transporter; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Complete proteome.
 SQ SEQUENCE 592 AA; 66708 MW; 3BFAB0EA350B9B0 CRC64;

Alignment Scores:

Pred. No.:	8,22e-69	Length:	592
Score:	1399.00	Matches:	277
Percent Similarity:	67.11%	Conservative:	123
Best Local Similarity:	46.48%	Mismatches:	178
Query Match:	32.89%	Indels:	18
DB:	16	Gaps:	6

AF133659 (1-2345) x Q8Y258 (1-592)

```

QY 361 ATAAAGCAATGCTTTCTATGTCGCCCCAAAGACAGGCCAGATCTACGACCTAGATT 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 8 Ileagserleuleuprotiyleutrpalia-----Tyrlystrhrpval 22
QY 421 GCCATTTGGCTGGATTTTGGGTGGTGAAGCCAGTAATTTGGTTCCCTTCATG 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 23 Alaualaileuserpheuilealaialaalsnleuglyvalprometval 42
QY 481 TTTAATATGCTGTGACAGCCTCAACACAGATGTCGGAAACATGCTGAACCTGATG 540
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 43 Metlysaigleuileasplameasn-----Valserprothrsp 56
QY 541 GCACCAATATACAGTTGCAACCATGGCAACAGAGTTCGATGGCTATGGTATCAAGA 600
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 57 ---Provalgalaileuvalvalprovalgylleileuglytyrclyleuaurg 75
QY 601 GCTGAGCTGCTTTTATACGAAGTTGCAATGCAATTTGGGACGATGCCAGAAAT 660
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 76 Leuserthrserleupheserleuaurguileuipheserlyvalthrgluser 95
QY 661 TCAATCCGAGAAATGCCAAATCTTCCATCTTCACATCTTCAACACTGATCTGGGTTT 720
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 96 Servalargthleuvalleuvalphearghlsleuhsalaileuserleuargphe 115
QY 721 CACCTGACAGACAGACGGAGCTTTATCTAAGCTATTCAGACAGAAACAGGGGTATC 780
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 116 Hlsleugluarglthrcllygmetsersargaplllegludrglythrargglyle 135
QY 781 ACTTTTCTGAGTCTTGGTATTAATCTTTCCATCATGTTTGAAGTATGCTT 840
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 136 GlnserleuileserlyserleuylserilleuProthrleuvalgluvalglyleu 155
QY 841 GTCAGTGGTCTTTGATATCAAAACGGGTGCCAGTTGCTTGGTGAACCTTGGAAAC 900
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 156 ValilethrtyrPhephevallystyrasphalaatrphalaileuilethrhecysala 175
QY 901 CTGTGATACATACAGATTCACAGTTGACAGTCACAGCGGTGAGAACTAGATTGAATA 960
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 176 Leuvalserlyrilevalpetherivalthrvalthrasttrpargthrtisshetrarg 195
QY 961 GAAATGAACAACAGATATATAGTAGAGTAATGCTGCTATAGACTCACTGCTGAATAT 1020
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 196 Argmetasngluleuaspserarglaalaenglnlyalaileaspserserleuasnph 215
QY 1021 GAAATGTGAAGTATTTAATAAGAAAGATATGAAGCAGACAGATATAGATTTTGG 1080
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 216 GlutrrivallystyrPhecllyasnngluylurlyglutthArargglyrpspdlubasnle 235
QY 1081 AAGAGTATGAAGTCTTCAATTTGAAGTAAGTACTCTGCTATGCTGAACCTTGGT 1140

```

```

DB 236 Arglystyrarglaalaalaileargserclnhiiserleuserleuasnphcgly 255
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1141 CAAAGTGTATTTTCAGNGTCGTTTAAACAGCTTAATGGTCTGCCAGACAGGAAT 1200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 256 Glnqlnleuilevalalaalaileuileuileuileuylrarglaaltnhnglyval 275
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1201 GTGCGAGGTACCTTACTGTGTGAGATCTAGTAATGGTGAATGAGCTGCTTTTCAGCTT 1260
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 276 Alaalaaglyhsmetthrleuglylasphevalleuvalaenthrlneumetleuglnle 295
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1261 TCATTTACCTTGAACCTTTCTGGGAACCTGTATATGAGAGCTAGACAAACACTATGAT 1320
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 296 TyrlleproleuasnphleuglyvaliletyrarggluileuylglnalavalthrAsp 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1321 ATGAACACCTTGTACTCTCAAGTGAACGCCAAATTAAGCAAAAGTATGAGCA 1380
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 316 MetasphargmetPheylsleuileuhisthrasnarggluvalalaasphargproasphla 335
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1381 TCTCCCTTCAGATCACACACAGACAGTACCTGGCCTTTGATTAATGTGCATTTTGA 1440
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 336 GlnProleuvalaVal-----ArgalaglygluvalArpheaalailevalAspPhegly 353
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1441 TACATTGAGGGCCAGAAAGTCTTATGGAATATCTTTGAAGTCCCTGCAGAAAGAA 1500
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 354 TyrgluserasnarglnlleuPheasphalaspPheThrileProalaglyThrThr 373
QY 1501 GTGCGCATTTAGAGAGGTAGTGGTCAGGGAAAGCAATAGAGAGCTATTTTTCG 1560
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 374 ThrAlaValValglylnsercllysercllysserthrleuAlaargleuileuphearg 393
QY 1561 TTTATGAGCCTCAAAAGGTAGACATTATCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1620
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 394 PheTyrsphalathrnsrcllyAlaileglnleaspglyglnasphalargAspvalThr 413
QY 1621 CTGAAAGCCTTCGAGGGCAGTGGAGTGTACTCAGATGCTGCTCTTCCATAT 1680
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 414 GlnlaservalarglaalaileglyylevalProglinsprrhvalleuPheasphasp 433
QY 1681 ACTATTTTATCAACCTTTATATGGAACATCAGTGGTCACTGACCTGAGAGTATGATCA 1740
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 434 SerilelytyrAsnillealaityrcllyArpProasphalathrarggluvalillela 453
QY 1741 GTGCGAAATATAGCTGTCGCTCATGATGCAATCTTTCGATGACATGATATGACACC 1800
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 454 AlaalaarglaalaileglnllehiscglyPhevalgluserleuProgluyltyrAspThr 473
QY 1801 CAAGTAGGGGAAACAGACCTCAACCTTTTCAGAGAGAGAAAGCAAGATGCAATTGCA 1860
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 474 Provalglylurargglyleuylleusercllygluylusglnargvalalailela 493
QY 1861 AGAGCAATTTTGAAGAGACCCCCAGCTCACTCATATGATGAAGTACTCATGCTTGAAT 1920
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 494 ArgthrleuileulysargproProilleuvalPheasphglnalathrSerlaileuasn 513
QY 1921 TCGATTTACTGAAGACATATCTTGTGTCGCATGAAGATGTCGTCGTCGTCGTCGTCGTCGTC 1980
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 514 SerargthrghlunhsalaileglncluleuunetargleuAlaaglnasnhilethrThr 533
QY 1981 ATTTTCATGACACACAGATTTGCAACAGTGGTATATGACATGAATCATGTTGCTGGAT 2040
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 534 Leuvalillealaahsargleuserthrillevalglyalaahsnglnlleleuvalmetclu 553
QY 2041 CAGGTATAGTACCCGAAACGGTATACCCAGCATGTTGCTGTCGTCGTCGTCGTCGTCGTCGTC 2100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 554 HlsglyargylleleuglurarglythrthlsalaserleuLeu---ArgalaglylurArg 572
QY 2101 TATTCAGAAATGTGGCATACACAGAGC-----AGCCGTGTGCAG 2139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 573 TyrlaaglnmettrpArgmetclnalalargglurProgluargvalgln 588

```

RESULT 13

Q984W3


```

Db 503 LysGlnArgValAlaIleAlaIleArgThrIleLeuYsaIaProProlIleLeuMetLeuAsp 522
Qy 1900 GAAGTACTTCTTCCTGTTGATTCGATTTACTGAGAGACTATTCTGGTCCATGAGAT 1959
Db 523 GlnIatTrSerIaIleAlaAspSerHisThrGlnGlnGlnIleGlnAlaIleAspLeu 542
Qy 1960 GTGGTCAACAGACAGACTCTTCTTTCTTTCATTCGACAGATTGTCAACAGTGGTTCATCA 2019
Db 543 ValSerIysGlnArgThrIleValIleAlaHisIleArgLeuSerThrValIleSerIa 562
Qy 2020 GATGAATCATTTGTTGGATCAGGGTAAAGGTAGCCGAACGTGTGATCCACCATGTTG 2079
Db 563 AspGlnIleIleValIleLeuLeuAspGlyGlnIleAlaGlnArgGlyThrHisValGlnLeu 582
Qy 2080 CTGTGTAACCTCATAGATCTATTCACAGAAATGTGGCATACACAGACGCCGTGTGAG 2139
Db 583 Met--ArgLysHisGlyLeuLysIleThrIleSerMetTrpAspArgGln----- 596
Qy 2140 AACCATGATTAACCCCAATGGAGCAAGAAAGAAATATATCCAAAGAGAGAGAAAGA 2199
Db 597 -----ArgGlnIatThrGlnIatGlnIatArg 605
Qy 2200 AAGAACTACACAGAAAGA 2217
Db 606 LeuArgLeuAlaIleArgGln 611

RESULT 14
Q8UGH3 ID Q8UGH3 PRELIMINARY; PRT; 629 AA.
AC Q8UGH3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ABC transporter, nucleotide binding/ATPase protein.
GN ATU1064 OR AGR C.1966.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
CX NCBI_TaxID=176299;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21608550; PubMed=11743193;
RX Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chien Y., Paulsen I.T., Eissen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RT Science 294:2317-2323(2001).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ourcillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Woliam C., Allinger M., Doughy D., Scott C., Lappas C., Makelz B.,
RA Flanagan C., Crowell C., Gurnson J., Lomo C., Sear C., Strub G.,
RA Cleio C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RT Science 294:2323-2328(2001).
RL EMBL; A0009070; AAL42077.1; ALT_INIT.
DR EMBL; A0008036; AAK6873.1; -.
KW Complete proteome.
SQ SEQUENCE 629 AA; 70102 MW; A957CB1E57061DAE CRC64;

```

```

Alignment Scores:
Pred. No.: 3,82e-87 Length: 629
Score: 1375.00 Matches: 281
Percent Similarity: 65.36% Conservative: 119
Best Local Similarity: 45.92% Mismatches: 190
Query Match: 32.32% Indels: 22
DB: 16 Gaps: 4

AF133659 (1-2345) x Q8UGH3 (1-629)
Qy 379 TATGTGTGGCCCAAGACAGCCAGATCTACGAGTACGATTCGATTCGCTGGATTT 438
Db TyrMetTrpProGlnGlyArgTrpAspLeuLysMetArgValIleTrpAlaThrValPhe 45
Qy 439 TTGGGTGGTGAAGGCCATGATATATTGTGGTTCCTTCATGTTTAATATAGCTGAGAC 498
Db LeuIleValAlaLysLeuValIleuIleAlaValProLysPhePheLysTrpAlaThrAsp 65
Qy 499 AGCCTCAACAGATGTGGGAAACATGCTGAACCTGAGTGATGACCAATATACATTGCA 558
Db AlaLeuAsn-----GlyLysLeuAspMetAlaGlyLeuValProAlaPheLeu 81
Qy 559 ACCATGGCAACAGCATTTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 618
Db LeuGlyAlaValAlaLeuValIleAlaIleAsnLeuThrArgLeuIleGlnValGlyLeu 101
Qy 619 AAGCAAGTTGGAATGAGATATTGGCAGGTAGCCAGATTCATCCGAATGAGATGCC 678
Db AsnGlnLeuArgAspSerLeuPheAlaSerValGlyGlnHisAlaValArgGlnLeuAla 121
Qy 679 AAAAATGCTCTTCATCTTCACACCTGATGCTGCTTTTTCACCTGACAGACAGACG 738
Db TyrArgThrPheValHisMetHisArgLeuSerLeuArgPheHisLeuGlnLysThr 141
Qy 739 GGAGCTTATATTAAGCTATTGACAGACAGAAAGGGGTATCACTTTTCCTGAGTCT 798
Db GlyGlyLeuSerArgValIleGlnArgGlyThrIleGlyIleGlnThrIleValArgPhe 161
Qy 799 TTGGATTTTATCTTCTCCATCATGCTGATGTTGAAGTGAAGTGCAGAGGTTTGCAT 858
Db ThrIleLeuAsnThrAlaProThrPheIleGlnPheLeuMetThrAlaIleIlePheTrp 181
Qy 859 TACAATGCGGTGCCAGTTTGTGTTGTAACCTTGAACCTTGACATTCATACACAGCA 918
Db AlaSerTyrGlyPheSerTyrValIleValThrValValThrValTrpAlaTrpIleTrp 201
Qy 919 TTCACAGTTGACATCACACGCTGAGAACTAGATTGATGAAATGAAACAAAGCAGAT 978
Db PheThrValArgAlaSerAsnTrpArgIleGlyIleArgAlaMetAsnAspSerAsp 221
Qy 979 AATGATGAGGTATGCTGCTATGACTGACTGCTGATTAATGAACTGTAATGATTTT 1038
Db ThrAspAlaAsnThrLysAlaIleAspSerLeuLeuAsnPheGlnThrValLysTrpPhe 241
Qy 1039 AATATGAAGAATATGAGACACAGACATATGATGATTTTGAAGCATATGAGATGCT 1098
Db GlyAsnGlnGlnMetGlnAlaIleArgPheAspValAlaMetGlnArgLysGlnLysSer 261
Qy 1099 TCATTGAAAGTACTCTACTCTGCTATGCTGAACCTTGGTCAAGTCTATTATTTCAGT 1158
Db AlaIleSerIleTrpTrpSerLeuGlyTrpLeuAsnPheGlyGlnGlyValIlePheGly 281
Qy 1159 GTGCGTTTAAGCATATATATGCTGCTGCGCAGTCAAGGAATGTGGCAGAGTACCTTACT 1218
Db IlegIysTrpIleIleMetMetValMetSerAlaLeuAlaValGlnArgGlyGlnGlnThr 301
Qy 1219 GTTGAAGATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278
Db IlegIysPhePheValPheValAlaAlaLeuLeuLeuGlnLeuSerValProLeuAsnPhe 321
Qy 1279 CTGGCAACTGTATATGAGACATGACACAGCACTCATAGATATGAAACACTGTTTACT 1338
Db IlegIysPheValTyrArgGlnIleArgGlnIleGlnLysThrAspIleGlnGlnMetPheAsp 341

```

QY 1339 CTACTGAGGTAGACACCCAAATTAAAGACAAAGATGATGCTCCCTTCAGATCACA 1338
 Db 342 LeuLeuGluValIGluValGluValThrAspArgProAspAlaValSPProLeuAlaValGly 361
 QY 1339 CCACAGACAGCTACCGTGGCTTGTATATATGCTATTTGTAATTCATTGAGGGCCAGAAA 1458
 Db 362 Pro-----GlyAlaIleSerPheArgValHisPheAlaTyAspProGluAspPro 379
 QY 1459 GTCCCTAGTGAATATCTTTGAAGTCCCTGAGAAAGAAAGGGCCATTGTAGAGCT 1518
 Db 380 IleLeuysGlyValSerPheAspValProAlaGlyThrValAlaIleValGlyPro 399
 QY 1519 AGTGGTGAAGGAAAGACAAATAGTGAAGGCTATATTTCTTATGAGCCCTCAAAAG 1578
 Db 400 SerGlyAlaGlySerThrIleSerArgLeuLeuTyArgPheTyAspIleGlnGlu 419
 QY 1579 GTTACGATTTATCTTGTGGTCAAAATATACAAATGTGAGCCCTGGAAAGCTTGGAGG 1638
 Db 420 GlyAlaValThrIleAspGlyGlnAspIleArgAspValThrGlnTySerLeuAspArgSer 439
 QY 1639 GCAGTGGAGAGTACCTCAGATGCTGCTCTCCATATACATTTATTTACAACTC 1698
 Db 440 MetIleGlyMetValProGlnAspThrValIleuPheAsnAspThrLeuAlaTyAsnIle 459
 QY 1699 TTATATGAAACATCAGTGTCTTCACTGAGAGAGTGTATGAGTGGCAAAATTAAGCTGGA 1758
 Db 460 ArgTyArgIleArgProSerAlaThrAspGluGluLeuysAlaAlaAspAlaValGln 479
 QY 1759 CTTGATGATGCAATTTCTTGAATGCGACATGATATGACACCCCAAGTGAAGGAAAGAGA 1818
 Db 480 IleSerAlaPheIleGlyIleuPheProAspGlyTyAlaThrMetValGlyGluArgGly 499
 QY 1819 CTGAGCTTTAGAGAGGAAAGAAAGAGTGAATGGAATGGAAGCCATTTGAAGAC 1878
 Db 500 LeuysLeuSerGlyGlyGlnGlnArgValAlaIleAlaArgThrIleLeuysAla 519
 QY 1879 CCCCAGTCACTCTATGATGAGAGCTACTCATGCTTGAATGCTGATTAAGAGACT 1938
 Db 520 ProIleuIleLeuIleLeuAspGluAlaThrSerAlaLeuAspThrIleThrGlnGln 539
 QY 1939 ATTCTGTGTCATGAGATGATGATGCTCAACAGAACTTATTTTCAATGACACAGA 1998
 Db 540 IleGlnSerAlaLeuAspIleValSerTyAsnArgThrThrLeuValIleAlaHisArg 559
 QY 1999 TTGTCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2058
 Db 560 LeuSerThrValIleHisAlaAspGluIleIleValLeuysGlnGlyLeuIleAlaGln 579
 QY 2059 CGTGTACCCACCATGTTGCTTGTCTTCACTCATAGTATCTATTTCAAGAAATGTCAT 2118
 Db 580 ArgGlyThrHisAlaSerLeuMetAlaHis---AspGlyLeuTyAlaSerMetTyPheSer 598
 QY 2119 ACACAGACAGCCGTGTGACAGAACCATGATTAACCCCAATGGAGAAAGAAAT 2178
 Db 599 ArgGln-----ArgGluAla 603
 QY 2179 ATATCCAAAGAGAGAAAGAAAGAAACTACAGAA 2214
 Db 604 IleArgAlaGlnGluMetLeuArgHisValArgGln 615
 RESULT 15
 Q92R12 PRELIMINARY; PRT; 627 AA.
 AC Q92R12;
 DT 01-DEC-2001 (TREMBLER). 19, Created)
 DT 01-DEC-2001 (TREMBLER). 19, Last sequence update)
 DE Probable ABC transporter ATP-binding transmembrane protein.
 GN R01120 OR SW000550.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Maes D.,
 RA Pohl T., Portecelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thepaut P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
 DR EMBL; AL591786; CAC45699.1; -
 DR InterPro; IPR001140; ABCcrampTM.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00654; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 627 AA; 70055 MW; 3C0BEA68107E451E CRC64;
 Alignment Scores:
 Pred. No.: 9.95e-87 Length: 627
 Score: 1369.00 Matches: 280
 Percent Similarity: 65.04% Conservative: 120
 Best Local Similarity: 45.53% Mismatches: 187
 Query Match: 32.18% Indels: 28
 DB: 16 Gaps: 5
 AF133659 (1-2345) x Q92R12 (1-627)
 QY 379 TATGTGCGCCCAAGACAGCCAGATCTAGAGCTGATGCTGATTTGCCATTTGCCGATTT 438
 Db 24 TyMetTrpProAlaAspArgIleAspLeuysLeuArgValAlaTrpAlaThrValIle 43
 QY 439 TTGGGTGTGCAAAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
 Db 44 LeuValIleAlaValAlaValIleValIleValIleValIleValIleValIleValIleValIle 63
 QY 499 AGCCTCAACAGATGTGGGAAACATGCTGAACCTGATGATGACCAATACAGTT--- 555
 Db 64 AlaleuAsn-----AsnIleProAspAlaLeuGlyPheLeuPro 76
 QY 556 -----GAAACATGCAACAGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
 Db 77 GlnPheLeuThrGlyAlaValMetLeuValIleuAlaTyAsnLeuAlaGlyLeuGln 96
 QY 610 GCTTTTATTAACGAAGTTCGAATGACGATTTGGCAAGGATGACCGAATTAACGCGA 669
 Db 97 AlaGlyLeuAsnGlnLeuAspAlaLeuPheAlaSerValGlyGlnHisAlaValArg 116
 QY 670 AGAATAGCCAAATAATGCTTTCTTCATCTTCAACACCTGATGCTGGTTTCACTGAGC 729
 Db 117 GlnLeuAlaTyLysThrPheValHisMetHisGlnLeuSerLeuAspPheHisLeuGln 136
 QY 730 AGACAGCGGAGCTTTATGTAAGCTATTTGACAGAGAAACAGGGTATTCAGTTTGTG 789
 Db 137 ArgArgThrGlyGlyLeuSerArgIleIleGluArgGlyThrTyArgGlyIleGluThrIle 156
 QY 790 CTGAGTCTTTGGATTTATCTTTCCATCTTCAACACCTGATGCTGGTTTCACTGAGG 849
 Db 157 ValArgPheThrIleLeuAsnSerValProThrLeuIleGluPheLeuThrAlaVal 176
 QY 850 GTTTTATTAACAAATGCGGTGCGAGTTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
 Db 177 IlePheTrpTrpGlyTyArgIlePheThrTyLeuLeuValThrAlaValThrValTrpLeu 196
 QY 910 TACACAGCATTCACAGTGTGACAGTCAACGCGTGGAGAACTGATTTAGATTAAGAAATGAC 969
 Db 197 TyIleTrpPheThrValArgAlaSerAspTrpArgIleAlaIleArgArgSerMetAsn 216

